HUMAN BIOLOGY
a record of research

SEPTEMBER, 1957

VOL. 29 No. 3

THE STATUS OF PHYSICAL TYPES

BY W. H. HAMMOND
Ministry of Health, London

THE present paper is concerned firstly with the general problem of body types as a form of classification and secondly with a comparison of two present-day type systems; somatotypes and factorial types.

It is necessary to consider the general question of types in relation to classifying individuals since many of the criticisms of somatotypes and factorial types could really be directed against any form of classification.

In the first place we have to employ some form of classification of human beings as also of human attributes, in order to understand them and even more to communicate our knowledge. We group particular items of conduct under labels—aggressiveness, anger, and so on—knowing that different individuals' manifestations of any emotion will differ slightly but will nevertheless be more alike than those of any emotion to which we give another name, and similarly physical characteristics such as chest girth have somewhat different constituent arcs in different people, yet we treat them as though they were uniform. In statistical terms the within-group variation (*i. e.*, trait) is less than the between-group variation. Just as a pattern of activity units, when

sufficiently coherent, constitutes a trait, so the traits shown by persons can be associated into wider and wider patterns finally covering the whole person when they become known as types. The type, therefore, is also the recognized pattern of characteristics which is rarely if ever the same for two persons yet which nevertheless has such family resemblance between certain individuals and not others that we can speak of the individuals as belonging to the same or different types.

The problem of conformity is the central one of type study and this is necessarily so, since types imply degrees of variation of their individual characteristics and not discreet changes as in species, and also because the characteristics themselves are not in any way unique. The usual methods of overcoming these difficulties are either to set limits within which the majority of characteristics must lie in order to belong to the type or else to regard the types as hypothetical pure patterns to which individuals approximate more or less closely. We then recognize inter-types in the one case or mixed types in the other.

Pure examples of types may be so rare as to be hypothetical and, therefore, the validity of a type classification should not be judged by whether individuals are distributed in discontinuous clusters round the type value or not but by the way in which they give logical coherence to a wide variety of aspects in individuals.

Although pure or extreme types are rare it is easier to understand types if they are described in terms of these extremes. Thorndike ('13) however is emphatically against this and claims that there is only one type, the mediocre and, in the sense that this is the most frequently occurring condition, this is true. Nevertheless, we are accustomed to referring to qualities by their extremes, as, for example, tallness and shortness, the average having no name. However, a more serious disadvantage is that when a complex set of qualities contributes to a type it is quite difficult to say which characteristics can be grouped into one continuum. This indeed is a major feature of types, for although tallness and shortness are opposites in height it is unusual to refer to tall or short types. Types imply different aspects which are not invariable or logical opposites nor yet completely related and they usually embrace a wide sphere, the whole of which may not be made explicit, much less be measured. It is relatively simple to classify persons in each of a variety of measurements separately, indicating whether a person is above or below the average and by how much. The difficulty is that there are so many possible measurements that

to describe a person in terms of some of them only gives a very incomplete picture—or if many are used the description becomes hopelessly unwieldy. Even if assignment of a plus or minus sign to some 50 different body measurements could give a fairly adequate description of an individual's physique, there would be 2^{50} classes if all combinations were possible enough to provide a unique class for every individual of the human race. The procedure would, however, have no value, for the usefulness of any classification is in proportion to the reduction of categories (consistent with homogeneity within them) required to characterize the maximum number of individuals in the largest number of aspects. Moreover, merely to add separate values is not very useful since two scores might be the same yet be made up in entirely different ways and the individuals having them might be totally unlike each other.

We obviously need to make a selection of characteristics which are in some way representative of a much wider number, and Type is the term for such a collection of qualities when it affects the whole of the individual or at least most of his physical or mental attributes. If we could obtain a limited number of qualities each having wider representativeness without overlapping they would be ideal to use in classifying individuals. 'Representativeness' is really the quality of forming part of a whole or group and to measure it we need to use one of the accepted methods of studying relations between qualities.

In physical measurements we usually acept the natural boundaries—arms, legs, trunk, and so on—and use one or two measurements out of an infinity of possible ones to represent their major dimensions of length, breadth or girth. These measurements are then correlated or co-variated, so determining their wider groupings, and the individuals who excel in the established groupings of measurements constitute the types. The selection can be further refined by employing multiple correlations which not only take account of the measurements which have the highest individual correlations with other aspects but also the maximum combined relation (through allowing for their own duplication with each other). These fairly straightforward methods at least should be applied to the study of the measurements chosen to measure types. Analysis of variance and discriminant function analysis offer more refined techniques. The chief difficulty of these methods is that they all require some outside criterion of the types and to this extent they involve a circular argument. Factor analysis is a device for determining the broad groupings of characteristics, whilst at the same time giving a

method of determining the relations between individual variables and these groupings. Theoretically, therefore, this appears to be a most promising approach to type study. Alternatively, we can compare and group individuals directly according to the pattern of their measurements, *e.g.*, according to whether their girth measurements are large or small relative to their own length and so on (in Burt's ('37) phrase we correlate traits or persons, or in American terms we use direct or inverse analyses). In one case we classify together the individuals who have small inter-individual variations in their measurements and in the other case we group the individuals whose intra-individual variation of measurements is similar. However, the process of classifying may not be explicit in this way nor be based on objective measurements; it may follow the more traditional methods of observation and subjective appraisal by general impression.

A brief historical survey will show how the methods of type study have developed. The first recorded use of types is that of Hippocrates who, over 2,000 years ago, observed a particular body form to be associated with a tendency towards lung diseases, the *phthisic* habitus and opposed to it the *apoplectic* habitus with its own temperament and disease liability. Similar types have been shown in the present century in Draper's gall bladder and ulcer constitutions. Humoral types then flourished in the Middle Ages when methods of study of human beings consisted more of philosophical constructions than that of practice field studies. The melancholic, sanguine, phlegmatic and choleric types were supposedly due to the presence of "black bile," blood, phlegm and yellow bile, in some respects foreshadowing the present endocrine disease groups due to over- or under-production of hormones. The purpose of these theoretical concepts was rather to organize people's understanding than to increase the sphere of knowledge, which is the principal aim of present-day experimental techniques.

Since these types were essentially personal systematizations of broad similarities between characteristics which were in themselves generally known, it is not surprising that different type classifications should arise during the centuries, depending upon the extent of the observer's penetration of thought and also on the interests of the times.

In the early 19th century a further resurrection of types occurred in the French *digestif*, *musculaire*, and *cerebral* types to which the type *respiratoire* was later added, and a somewhat similar manifestation ocurred in faculty psychology.

After a period during which mental faculties came into disrepute and physical types were little heard of outside Italy a resurgence occurred in the field of anthropology in the form of race types. At the same time, in psychology Jung's introverts and extroverts and in psychiatry Krestchmer's pyknic, athletic and asthenic types re-awakened interest in type classifications and for the first time they brought a quite new interest in their train, a stimulus to practical research work. These types, in common with all previous ones, represented hypothetically pure manifestations which may never be realized in any one individual. The pure types began to be used as reference frames for charting human beings who, in the case of race types, were classified as mixtures of pure ethnic groups, whilst the psychological types became tendencies possessed in different degrees by individuals.

The problem of measuring the type tendencies now occupied an entirely new importance and this placed a new emphasis on the elements constituting the type rather than on their wholeness. Thus different racial characteristics such as head or face shape were measured, and questionnaire assessments and, in some cases, tests were employed to measure general characteristics which were supposedly diagnostic of the personality types. At the same time the types themselves were characterized more rigorously than was possible before. Kretschmer's ('25) studies gave a new synthetic approach in which physical types were associated with mental dispositions mainly, in the beginning, through the relations between extremes. The cyclic mental diseases (manic depressive psychoses) were associated with pyknic (rotund) build and schizophrenia with asthenic (lean-slight) build. This association with physical types was then extended to less extreme cases, cycloid and schizoid tendencies, and to the cyclothymes and schizothymes of the normal personality range.

The genesis of the two modern type systems, somatotypes and factor types can be related to the traditional types; Sheldon's ('40) Somatotypes, inspired largely by Kretschmer's types, are logical descendants of types as pure patterns. On the other hand factor types arose very differently. They sprang from Spearman's ('27) attempts to account for the varying relations between charactertistics in terms of factors or 'causes' which were common to them all or specific to any one. From the extension of this to include only part of the correlations of a table, followers of Spearman provided the basis of present day body types. In the beginning the nature of the physical types was incidental to the

use of body measurements as material for demonstrating the different factorial methods, the chief advantage of using physical measurements being that the mathematical constants of the analyses could readily be equated to physical attributes which would make them clear to non-mathematical readers and also suggest new techniques to the factorists themselves. Kretschmer's and Sheldon's types, therefore, arose from directly comparing individuals (in Kretschmer's case mental patients with different forms of insanity) and measuring their characteristics, whilst factor types were based originally on comparing characteristics and not persons. The chief difference between factor types and previous attempts at type study was that the factorists initially had no pre-conceived ideas of what types might exist; that is, the data were to demonstrate whatever groupings or clustering of characters might be present. However, it is true that more recently different methods of analysis have been developed and these have to some extent governed the kind of types which result. In particular they have determined whether the groupings of traits correspond to a co-ordinate classification (group factor analyses) or a super- and sub-ordinate classification (bipolar general factor analyses).

Having considered methods of type study in general we can now consider in greater detail the two main type systems of the present day, somatotypes and factor types, and show how far they are alike or different from each other.

Somatotypes.

Sheldon's somatotypes, named endomorphs, mesomorphs and ectomorphs from the predominance of structures derived from the three embryonic tissues, endoderm, mesoderm and ectoderm, respectively, are obtained from photographs taken in the nude from front, side and back with the body posed in standard positions. Originally the typing was done visually without reference to measurements and it was based in the beginning upon a subjective general impression similar to that by which the earlier classical types were postulated. Measuring techniques, which it is hoped will eventually replace the 'scopic' method, are being developed along with refinements of the rating procedures (especially by Tanner ('49) in England and Dupertuis ('50) in America).

The subjective nature of somatotypes has its disadvantages but certain advantages are claimed for it. A skilled observer can make use of the subtler distinctions of body outline which are not amenable to

measurement. For example, the body contours can show where a person who is now thin was once fatter and this might indicate (on the somatotyping basis which seeks to assess fundamental characteristics and not the momentary conformation of measurements) that the true rating would be more endomorphic and less ectomorphic than would be inferred from measurements alone. The characteristics of the somatotypes are considered to be all embracing so that no physical characteristic can be neutral to or outside this continuum. A person's hip, arm, chest, etc., are characteristic of one or other of the three types and hence contribute to the score for that type. They are never considered to be without effect and the most can happen is for an individual to be dysplastic or to have aspects which follow one type in one part of the body and a different type in another part, *e.g.*, shoulder and trunk may be endomorphic and legs or arms ectomorphic. The three components were recognized as inadequate to cover all aspects of physique, hence another modality, the degree of bisexuality or gynandromorphy which can affect physique independently of the three somatotypes was also introduced.

There is no mention of general size since the 'scopic' method is concerned with relative shape and the measurements are always expressed in relation to height.

The system is not yet applicable to children because standardized rating procedures are not available for these ages.

A high association is claimed between endomorph, mesomorph, and ectomorph body form respectively and the three temperamental types, viscerotonia (relaxed, comfort-loving, tolerant, needing affection and others' company), somatotonia (assertive, physically active, direct in manner, callous, noisy), cerebrotonia (restrained, taut, liking privacy—especially when in difficulty, unsociable and having unpredictable attitudes). The system of scoring for the temperamental types is similar to that for somatotypes, Sheldon ('42).

The main defect of Sheldon's system is that the somatotypes are basically subjectively determined and no measurement procedures will alter this unless the types themselves are modified as a result. The method of equal-appearing intervals should obviously be abandoned when the elements constituting the types are made explicit. A second related difficulty is that the scale intervals 1-7 (or half units if these are employed) may not represent additive units as in an extensive magnitude and this limits statistical manipulation. This is especially true since the three components are also inter-related in a complex way. The

distribution of values 1-7 in each component could ordinarily be studied to see whether 4 really represents the mean or not and how far the limitation of the extremes to 1 and 7 may be truncating the distribution, but when these are also based on values found in the other two components this becomes impossible. We can not say whether a particular combination, *e.g.*, a 7 for endomorphy and 3 for ectomorphy does not exist because it is logically impossible in the same way that a 7 for tallness and a 3 for shortness would be impossible or because such an example has not yet been found. Sheldon resorts to an arbitrary limitation: that the components must not total more than 12. The fact that there is a higher negative correlation between endomorphy and ectomorphy than between any other pairing of components is not taken into account.

Logically the somatotyping system and its measurements are somewhat unsatisfactory. To select as components three particular combinations out of a number of complexly inter-related physical qualities without knowing how the separate characteristics contribute to them nor how the components are related, is rather like attempting to give a point's position in a spherical space in terms of the ribs of an umbrella of unknown length set at unknown angles to each other. Obviously, the situation needs tidying up. This can only be done satisfactorily (a) when the separate physical characteristics are measured, (b) their relation studied to see whether they cohere to form groups which correspond to the labels endomorph, mesomorph and ectomorph, (c) the relations between the type groups are studied, and finally when (d) units of type measurement having meaningful intervals are devised, *e.g.*, a difference between 3 and 4 for ectomorphy should be in some way equivalent to a difference between 5 and 6, if only that it selects a similar number of individuals from the total distribution.

The interpretation of types as derived from the three embryological tissues is dubious, particularly in the case of ectomorphs, since their ectodermally derived tissues are not really better developed than in the other types, but it is only that the endodermal and mesodermal structures are less well developed. This is only a secondary criticism of somatotypes, however, and the names could be regarded merely as convenient rather than as indicating a causal basis of the somatotypes.

Factorial types.

There is really no single set of factor types comparable with the

somatotypes. Many different workers have set out independently to show the presence of body types from various batteries of measurements using different factorial procedures. (Burt, '37, '44, '47; Cohen, '39; Hammond, '42, '53a and b, '57; Eysenck and Rees, '45; Banks, '47 in England and Carter and Kraus, '36; McCloy, '40; Mullen, '40, Moore and Hsu, '46; Thurstone, '46; Howells, '49, '52; Sills, '50; Heath, '52; and Lorr and Fields, '54 in America.) The basic approach has usually been to calculate correlations or co-variances between measurements and analyze them in terms of general influences or factors which are regarded as accounting for (a) all the correlations or co-variances in the battery (general factors), (b) some of the correlations only (group factors) often superimposed upon a general factor, and (c) unique influences (error or specific factors) which only contribute to the separate measurements and so reduce the correlations with other variables. Any combination of the above factors may be used but the types are usually based on the group factors, direct bipolar contrast factors or on the rotated general factors of simple structure rather than on the first general factor of a centroid analysis. The types then comprise those individuals who excel in any one of the groups of measurements. The factors corresponding to the groupings of measurements are usually independent (in England), but oblique factors or transformed secondary factors are used extensively in America. Having found which measurements go with which the next operation is to measure the strength of the factors so as to give them type scores. These are usually calculated from regression equations for the different factors in terms of the measurements, that is, as weighted sums of measurements of a particular kind. In the case of contrasting types, measurements may give a minus contribution to the total, *e. g.*, ectomorphs or leptosomes will have high values for length growth and low values for girths and conversely for endomorphs or pyknics.

The exception to the above procedure is to correlate persons' patterns of measurements directly with each other and then analyze them to provide groups of similar persons. The characteristics of the type groups, *i. e.*, their trait relations, are then studied. From these Burt ('37) derives a type standard against which any individual's measurements may be correlated to indicate his type conformity score. Howells ('52) determines a type factor measurement from the measurements weighted by the individual's factor saturations, that is, if individuals who have high saturations for the endomorph factor also have large

trunk girths, the girths are given high weight in determining the factor measurement, whereas structures in which the extreme types show little difference have little weight.

Factorial studies are as objective as any type investigations so far, the chief limitations being their dependence upon the initial choice of tests or persons studied and the fact that the solutions are not unique. However, the results from any factor analysis procedure can generally be translated into those of any other kind by suitable transformation matrices. The outcome of any analysis does not therefore simply depend upon the form of analysis used, otherwise it might be true to claim that the factorist merely gets out of the analysis what he puts into it.

Although there is no commonly accepted type factor solution equivalent to somatotypes, there is a high measure of agreement between analyses of various batteries of measurements by different investigators. Without exception the English studies find that the major distinction is between growth in length and either breadth or girth growth, depending on which measurements are included in the battery. This is true of European and non-European adults (Burt '44), normal and psychotic subjects (Cohen '39 and Eysenck '45) and also of children down to infancy (Hammond '57). Analyses of the same battery of measurements in different yearly age groups show very high consistency in the factor saturations for the different measurements (Burt and Banks, '47; Hammond '53).

Studies on American adults (McCloy, '40; Mullen, '40; Heath, '52) also confirm this distinction, which is also shown by the analysis of measurements for garments and pattern construction based on American children (Hammond '53a) and infants (Carter and Kraus '36).

Most of the batteries of measurements analyzed were taken from existing studies made for a variety of purposes. In the English studies the methods used necessarily contrasted girths, breadths or other measurements with the lengths. Burt, therefore, uses the non-committal term pachysome or bulky. Yet breadth and girth measurements are themselves distinct and it is necessary to determine whether the girths or breadths are the more fundamentally opposed to lengths, or whether these merely form an antagonistic group only in the sense that something must provide the negative aspect of lengths if these represent positive deviations about a mean. When lengths, breadths and girths are all present it is usual to find the girths more strongly opposed to length in the first factor of a bipolar analysis, after which lengths can

be split into trunk and limb lengths and the other body measurements into girths contrasted with breadths. In the case of group or oblique factors the girths again form the most prominent grouping after lengths. Perhaps the critical point is that fat measurements have extremely high saturations with the girth group whilst the breadth measurements appear to be related to girths mainly through their direct contribution to girth measurements taken at the same levels rather than because of any fundamental affinity.

Height, trunk girths and breadths are composite measurements dependent upon growth of different long bones, muscles or fat and Heath ('52) prefers to identify her secondary factors with development of these tissues rather than in terms of the composite measurements. This is a useful interpretation especially as fat measurements taken at different body sites have extremely high intercorrelations and similarly for the long bones. X-ray measurements of different muscles or measurements of girths, adjusted to allow for fat, both suggest a strong group factor for muscle development also.

The description of human beings in terms of factors has its disadvantages, although it approaches more closely to the general requirements of type systems as set out earlier. In the first place, any disadvantages which apply to correlations also apply to factor analysis of them. In the second place the method is based on the interpretation of correlations in terms of the ratio of common elements to total elements. Although this presents no difficulties when applied to single correlation coefficients, when several variables are interpreted in this way there are many different possible models which can be fitted to the observed correlations involving elements of different generality. For example, it is possible to account for a table of correlations in terms of separate groups, each of which overlaps, or in terms of a common factor. Mathematically, the effect is the same but the interpretation in terms of correlated types or a general influence and uncorrelated types gives a quite different emphasis. But if there is no unique factor solution to a given set of correlations at least the various assumptions involved in the different methods are usually made explicit and within these assumptions, they can be checked. However, the effects of sampling errors are not easy to estimate in order to test rigorously the goodness of fit of different solutions.

It is not a necessary condition of factor analysis that the factors produced should be anything more than principles of classification of

measurements as they exist at any moment for a particular group of subjects. But their maximum usefulness will only be realized if the factors can be identified with more permanent features. However, the identification of factors with outside causes can also lead to disagreement. In so far as these factors correspond to groupings of measurements from different studies there will naturally be more disagreement between factor types than is the case in somatotypes where the position is equivalent to accepting a particular battery of tests and the factor measurements derived from them as put forward by one investigator.[1]

Another difficulty is that, although the type groupings may be accepted, it is a rather complicated matter to assess the type strength of individuals, especially if the general growth has to be allowed for first.

The relation between somatotypes and factor types.

The methods of factor analysis can be applied to the measurements of somatotypes just as to any other measurements. Several studies have been made to attempt to validate the somatotypes. These consist of correlating the measurements of individuals who have previously been somatotyped by inverted analysis to see whether the grouping factors for persons agree with the somatotyping and then determining the composition of the measurements which differentiate them. The second method is to compare the factor analysis of measurements to see whether they are grouped into ways which are equivalent to the characteristics of the somatotypes. In the first kind of analysis Howells ('52) and Lorr and Fields ('54) agree remarkably closely in the person factors they found. Both agree in regarding the endomorphs and ectomorphs as forming one continuum, the first two oblique factors comparing closely in respect of the mean factor saturations for the somatotype groups (see table 1).

Factor I differentiates between ectomorphs and endomorphs in both cases; Factor II differentiates mesomorphs from endomorphs or ectomorphs. The division in Factor I is more towards the ectomorph end of the scale in Howells' analysis and his second factor compensates by being slightly nearer the endomorph pole.

[1] There is much to be said for a convention of factorists meeting with the object of recommending standard batteries of physical measurements for assessing varieties of physique. Such standardization would enable different workers who are interested primarily in practical problems to obtain type scores which will at once be simple and permit exact comparisons to be made with the results of others.

Sills ('50) correlated the somatotype ratings as variables along with a number of measurement ratios and also found that ectomorphy was the opposite pole of endomorphy rather than a separate factor. In this case, however, the subjects were not selected initially from extreme somatotypes. In general, the factor analyses fail to substantiate the somatotype groupings as the best solution to the measurement data. It is true that a centroid analysis with orthogonal factors could not possibly agree with Sheldon's threefold correlated somatotype groups. As Tanner ('52) has pointed out, if the ectomorphs coincide with Burt's

TABLE 1

Factorial analysis of somatotypes

	Factor I		Factor II	
	MEAN SATURATIONS		MEAN SATURATIONS	
	Howells	Lorr and Fields	Howells	Lorr and Fields
Endomorphs	.84	.80	—.08	—.20
Mesomorphs	.12	—.12	.35	.52
Ectomorphs	—.92	—.73	—.10	—.13

leptosome factor it should be impossible to find the equivalents of the other types among later orthogonal factors. But in fact, the pachysome pole of the first factor in some batteries does approximate very closely to the endomorphs. However, even where the analyses are based on specially selected extreme somatotypes and the form of analysis is such that the factors could correspond to Sheldon's types they do not agree closely. Two studies (McCloy, '40; and Sills, '50) claim to have isolated a factor omomorphy characterized especially by wide shoulders compared with hips more reminiscent of Kretschmer's athletic type. Howells ('52) concludes that "certain other kinds of differentiation come to light which are now masked by the recognition of endomorphy and ectomorphy as independent components bearing attributes which are not simply negatives of one another" and he likewise questions the concept of dysplasia in view of the "non-Sheldonian kinds of differentiation which are consistently and sensibly organized."

In passing, it seems to me to be a mistake to omit consideration of the general size or general shape factors by selecting individuals to minimize these. The general factor contributes so overwhelmingly to

the total variance especially in children, that it has by far the greatest predictive value for all measurements, and the method of comparing individual measurements relative to height as the means of allowing for it, is an unsatisfactory one.

Factor schemes, such as those described, are complicated and highly technical so that they may lose in usefulness what they gain in precision. It is worth while, therefore, to see what simple steps may be taken to produce a workable typing system, making use of the best features of the somatotypes and factor types. Parnell's ('52) use of standard deviations in different measures grouped into muscles, fat or 'leanness' groups is a good basis on which to begin.[2] If we substitute bone length measurements for height/weight (which is a composite of length and absence of muscle or fat) we shall then have three measures of tissues each of which will be equivalent since it is based on standard deviation units.[3] The individual's general size will be indicated by the sign and magnitude of these units. (At present in somatotyping one could have a large or small individual of any given type figure, and the figures give no indication of this). On the proposed scheme the relatively muscular, fat or long types would show in the pattern of the numbers whilst at the same time their 'size' would be indicated by the level of the numbers. For example, a $+7+7+7$ (a possible figure) would be a person who was among the biggest 2 per 10,000 in muscle, fat and length and yet he would be of average proportions—a sort of giant in fact. Similarly, a $-7-7-7$ would be a scaled down individual or dwarf, and $0 \cdot 0 \cdot 0 \cdot$ would be the individual corresponding to the average measurements in all respects. If we designate the variables, L. M. F. or length, muscularity and fatness, + will represent long, muscular or fat and — short, amuscular or thin. The various combinations would be long-muscular-thin type $(++-)$, short-muscular-fat $(-++)$ and so on. The terms large or small could be used to summarize the figures qualifying them with whatever component did not conform, *e.g.*, the first person above would be a large thin type, the second a large short type, 'large' because both were above average in the majority of measure-

[2] Hooton according to Hunt ('52) uses a similar scheme measuring fat, muscularity and 'attenuation' based on the ponderal index.

[3] Actually further precautions are necessary to allow for the extent to which the particular muscle, fat or long bone measurements are correlated with each other, and the measurements would need to be carefully chosen for each of them to indicate general size as well as its own sub-group.

ments. These may be rather unusual sounding designations but in any case verbal labels are unsatisfactory substitutes for figures as we discover at once when we come to describe the different possible combinations in verbal terms. It is because the dividing lines between the terms are ill-defined that labels should only be used as a very rough guide and, in any written work, numerical values should always be given along with the verbal descriptions to define them. Body photographs may still be used to study these types.[4]

The above scheme at least has the advantage of using combinations of measurements which conform to established group factors in factorial studies, whilst the calculation of the type scores is much easier than the regression method and at the same time makes a distinction between overall size and type strength. (The predominance of one component could be expressed as the difference between its type and the average, whilst shapeliness—not in the artistic sense—would be the sum of the differences, between paired components 1-2, 2-3, and 1-3.)

As no agreed set of criteria is available for somatotyping children it is especially useful and necessary to arrive at an agreed typing system in this field whilst the position is open. By so doing it would be possible to obtain the great advantage of unanimity which adult somatotyping offers without sacrificing the advantages in respect of objectivity and rigorous treatment offered by factor studies. Even in studying adults it would be well to test the validity of somatotypes by measurements in unselected populations rather than in groups of extreme somatotypes, for if the somatotyping grouping is not confirmed (and these are the indications from the available studies) then substituting measurements for the subjective appraisal of somatotypes would appear to be of dubious value. Another consideration is that children's types will need to be related to those of adults. Children's types have also been found to have a different basis from that of somatotypes, consequently continuity with present adult somatotypes would appear to be impossible.

Are types necessary at all?

So far we have been mainly concerned with the rationale of any assessment of multiple types in order to show how some of the more desirable features could be introduced and substituted for present prac-

[4] The types would represent the body conformation at a particular moment without any claim to genetic basis (but see p. 238).

tices. One might take the more devastating view that any kind of type study is too stultifying to warrant this attention to details. Do types, however well authenticated they may be, help us at all? Pure expression of the types is rarely met with and the majority of individuals are somewhere near the mean. To use type terms is somewhat like describing an average person as being compounded equally of extremes, *e.g.*, of tallness and shortness. Those whose study of types has taken place among extremes tend to regard the normal in these terms whilst others often consider the extremes as pathological conditions constituting distinct phenomena having no connection with the normal. I believe that factorial methods of type study can provide a truer evaluation of the position. They represent (or should represent) a necessary aspect of the statistical study of inter-related variables upon which the study of the complex field of human behavior must at present rely if it is to become a science. The use of trait rating scales as units in psychology may represent a transitional stage until more fundamental behavior units can be discovered. Similarly physical or mental types represent a further stage of generalization but both have their usefulness at different levels of behavior study. Thus if we want to know how a person is likely to perform under a given set of circumstances we can use these circumstances as actual test situations. However, if we know from type study that certain types exhibit a whole series of characteristics, then we need only test or observe a few of these characteristics to be able to predict the others. If in addition to establishing physical types we can say that these are associated with mental characteristics—forming psycho-physical types; then we can extend the process correspondingly—always assuming that the degree of relationship is sufficiently close to afford some improvement in prediction compared with the situation in which the physical types are not known. To test this we need to be able to specify the type adequately and then to study the widest possible influences of the type conformation.

One of the chief objects of such study should be to test how consistent the types are over a number of years, for unless the individual's type make-up has some constancy we can hardly expect it to have any predictive value in other fields. The author's own work on follow-up studies in factor types of children (Hammond '57) indicates that the same kind of types (from attained measurements) can be shown at all ages and growth can also be analyzed into similar type patterns. However, the type constancy is little more than we should expect if a child's

type were established very early in life and the subsequent growth were random or neutral to the type. That is, on average, past growth appears to add comparatively little to the predication of the type differentiation. There must necessarily be some type consistency in children because of the contribution which the earlier measurements make to the later ones and as childhood advances this becomes proportionately greater compared with the growth which is capable of upsetting the type. Correlations of an order up to 0.8 are therefore to be expected over a period of two years on this account alone and this is the order of correlation actually found.

There are also immediate objects of study of physical types, such as the relationship between anatomical characteristics and performance in different occupations or sports, in order to use the type as a basis of selection for recommending physical activities in which the individual is most likely to succeed. It may be that incentive is more important than physical build but physique may set limits to what is possible. If one may speculate on this point it is to urge the use of types not so much as an instrument of prediction in themselves, but rather as a preliminary to studying certain of their characteristics more intensively. For example, some physical types appear to be more susceptible to certain diseases than others, yet the relation may be too loose to aid prediction directly. However, the particular individuals of a type having the disease may be characterized by some special aspects of the type syndrome only and these might be made the basis of future prediction.

The main aims of type study should therefore be, firstly, to make the types and their measurements more objective and refined, secondly, to reach general agreement on these types so that further studies will be comparable, and thirdly, to devise simplified methods, such as those suggested, for measuring the type strength.

SUMMARY

The present paper adopts a historical approach to the study of body type and sets out some theoretical requirements for any classification systems and also the special requirements of type as principles of classification exemplified in the two current systems, somatotypes and factor types.

The chief weakness of somatotypes is that the types themselves are arbitrary and complexly related whilst the different measuring scales

have not been studied sufficiently in relation to each other and to the types, and the form of the measurements makes it impossible to construct theoretical distributions for these variables.

Factor types offer no generally agreed solution equivalent to the somatotypes with the advantages of comparability for different investigators, but they have a more objective and theoretically advanced basis although they too are dependent to some extent on the measurements selected and the factor method used. The major physical types agree closely at different ages but the complication of assessing factor types by regression equations presents a serious difficulty in popularizing factorial methods.

A compromise type procedure is proposed (especially for use with children where somatotyping is not yet possible). Three components, length, muscle and fat are measured in standard deviation units as in Parnell's scheme and expressed by positive and negative numbers whose magnitude indicates general size and whose pattern shows the type distinction. The scheme would be based on direct anthropometry supplemented by photographs were possible.

LITERATURE CITED

BURT, C. 1937 Correlations between persons. Brit. J. Psychol. *28*: 59-96.

——— 1944 The factorial study of physical types. Man *44*: 82-86.

——— 1947 Factor analysis and physical types. Psychometrika *12*: 171-188.

BURT, C. AND CHARLOTTE BANKS 1947 Factor analysis and body measurements for British adult males. Annals of Eugenics *13*: 238-256.

CARTER, H. D. AND RUTH H. KRAUS 1936 Physical proportions of the human infant. Child Develop. 7: 60-68.

COHEN, J. 1939 Physique and proportion. Brit. J. Med. Psychol. *18*: 323-337.

DRAPER, G. 1924 The human constitution. Philadelphia. Saunders.

DUPERTUIS, C. W. AND J. M. TANNER 1950 The pose of the subject for photogrammetric anthropometry. Amer. J. Phys. Anthrop. n. s. *8*: 27-48.

DYSENCK, H. J. AND W. L. REES 1945 A factorial study of some morphological and psychological aspects of human constitution. J. Mental Science *382*: 8-21.

HAMMOND, W. H. 1942 An application of Burt's multiple general factor analysis to the delineation of physical types. Man *42*: 4-11.

——— 1953 The determination of physical type in children. Human Biol. *25*: 65-80.

——— 1953b Body measurements of pre-school children. Brit. J. Prev. and Soc. Med. *9*(3): 152-158.

——— 1957 The anatomy of physical types as determined by factual analysis. Human Biol. *29*: ——.

HEATH, HELEN A. 1952 A factor analysis of women's measurements taken for garment and pattern construction. Psychometrika *17*: 87-95.

HOWELLS, W. W. 1949 Factors of human physique. Amer. J. Phys. Anthrop. n. s. *9*: 159-191.

——— 1952 A factorial study of constitutional type. Amer. J. Phys. Anthrop. n. s. *10*: 91-118.

HUNT, E. E., JR. 1952 Physical constitution. Amer. J. Phys. Anthrop. n. s. *10*: 55-73.

KUTSCHMER, E. 1925 Physique and character. Kegan Paul. London.

LORR, M. AND V. FIELDS 1954 A factorial study of body types. J. Clin. Psychol. *10*: 182-185.

MCCLOY, C. H. 1940 An analysis for multiple factors of physical growth at different age levels. Child Develop. *11*: 249-278.

MOORE, T. V. AND E. H. HSU 1946 A factorial analysis of anthropological measurements in psychotic patients. Human Biol. *18*: 133-157.

MULLEN, F. A. 1940 Factors in the growth of girls. Child Develop. *11*: 27-41.

PARNELL, W. J. 1952 Recording human constitution. Eugenics Review *44*: 20-27.

SHELDON, W. H. AND W. B. TUCKER 1940 The Variety of Human Physique. New York. Harper Bros.

SHELDON, W. H. AND S. S. STEVENS 1942 The Variety of Human Temperament. New York. Harper Bros.

SILLS, F. D. 1950 A factor analysis of somatotypes and their relationship to achievements in motor skills. Res. Quart. *21*: 424-437.

SPEARMAN, CHARLES 1927 The Abilities of Man. Macmillan. London.

TANNER, J. M. 1947 The morphological level of personality. Proc. Roy. Soc. Med. *40*: 301-308.

TANNER, J. M. AND J. S. WIENER 1949 The reliability of the photogrammetric method of anthropometry. Amer. J. Phys. Anthrop. n. s. 7: 145-186.

THORNDIKE, E. L. 1913 Educational Psychology. Vol. 3. New York. Teach. Coll. Columbia Univ.

THURSTONE, L. W. 1946 Factor analysis and body types. Psychometrika *11*: 15-21.

SEX DETERMINATION FROM THE SKELETON

BY FRED P. THIEME AND WILLIAM J. SCHULL
University of Michigan

SEX, unlike most phenotypic features in which man varies, is not continuously variable but is expressed in a clear bimodal distribution. This bimodality, or polymorphism, is determined and maintained by the inherent nature of man's genetic-reproductive mechanism. In this respect it is dissimilar to most other polymorphic traits, such as the allelic varieties of a blood type, which are primarily maintained by genetic equilibrium and/or the balance of selective mechanisms. But sexual dimorphism is similar to other expressions of polymorphism to the extent that clearly defined, genetically determined forms, or phases, of the same species co-exist in single breeding populations. Similarity also exists in terms of classification; in both sexual and other polymorphic variability clear categories exist. Ambiguous forms are non-existent or exceedingly rare since the phenotype is a clear expression of the genetic determinants.

In the problem of determining the sex of the skeleton, it is important to note that the phenotype of sex, while it is an expression of a distinct genotype, is fundamentally different from other kinds of polymorphism on two counts. As previously mentioned, it differs first as a consequence of its being determined by a special genetic-reproductive mechanism which is at the chromosome level and thus involves many genes; and secondly, as a consequence of this, sex differences are expressed in many aspects of the phenotype through widely varying levels of genetic influence. Thus, some features of the phenotype give clear indications of sex while others may be of little or no use.

The extent to which the skeleton, as one part of the gross phenotype, can be employed in determining sex is of interest in physical anthropology. From what is known, it seems clear that three varieties of

influences cause sex differences in the skeleton. The first is reproductive function which is primarily expressed in variability in pelvic morphology; secondly, there are genetically influenced contrasts in gross bone size and proportions; and lastly, there are differences in the amount and proportion of body musculature as shown by attachment markings on the bone. These various influences are not always separable in their effects but, in the main, they cause the phenotypic differences in skeletal morphology which demarcate sex.

The determination of sex of skeletons has always been of importance to anthropology. In most archeological investigations a classification by sex is the first step taken in the analysis, and important conclusions about the nature of the group and their culture are based on sex ratios. In studies of human evolution, important questions of relationships depend upon the accuracy with which sex is assigned. While effort is made constantly to improve the accuracy of sorting and the literature is full of references to "hallmarks" of sex which are said to be of value, few investigators know the level of accuracy that they achieve or even should expect from the method or methods they use. While there is no substitute for experience in such matters, it is essential that work be done to establish the levels of accuracy achieved by various methods of sexing.

This paper reports the results of an investigation designed to find the sex discriminating efficiency of several measurements on the postcranial skeleton, using specimens of known sex. The study was initiated in order to accurately assess the discriminating value of several easily measured features, rather than to find new "hallmarks" of sex in the skeleton. The selection of measurements was based upon the general considerations stated previously. Analysis of data was arranged so that a precise probability statement could be made about the usefulness of the measurements for sexing. The general plan, and summary of the investigation is designed so that it might have the highest practical value for archeology, physical anthropology, and forensic medicine.

The expectation is high that sex can be determined with accuracy if the specimen is adult and complete. If it is immature or incomplete we may be merely guessing with little assurance of being right. The results of this study indicate that the upper limit can exceed 95%. It is up to the experienced investigator to estimate, on the basis of completeness, age, and in light of the particular problem he faces, what the probability of the correct assignment of sex may be.

MATERIALS AND METHODS

Known adult male and female Negro skeletons from the Terry Collection were measured. They were made available for our study through the courtesy of Dr. Mildred Trotter and the Department of Anatomy of the Medical School of Washington University. Support for this study came from the Wenner-Gren Foundation for Anthropological Research to the Laboratory of Physical Anthropology at the University of Michigan. Measurements were taken on the left side of the 99 known male and 101 known female Negro skeletons. The measurements used in the discriminant analysis were as follows:

1. Femur Length (Bicondylar)
2. Femur Head Diameter
3. Humerus Length
4. Epicondylar Width of the Humerus
5. Clavicle Length
6. Ischium Length
7. Pubis Length.

Other measurements taken were: sternum width (measured between the estimated centers of the surfaces of clavicle-sternum articulation) and the lengths of the right clavicle, humerus, and femur. These latter measurements were taken so that right-left long-bone asymmetry could be examined and also shoulder width could be estimated. Inasmuch as females have relatively broader hips and narrower shoulders than males, it was thought that an ischio-clavicle index might have special value for sexing. This did not turn out to be the case as it was of less value than the ischium-pubis index (I/P). The pattern of right-left asymmetry also had no value for sexing, but since it may be of interest, the results are noted. The opening of the sciatic notch was also observed and recorded. Values were assigned according to the degree of opening with the number 1 recorded for the exaggerated male condition and 5 for the extremely open female notch.

All measurements were made by one individual, F.P.T., and the values were punched and verified on IBM cards. Using an IBM electronic calculator, we obtained values useful for comparison and for discriminant analysis. The discriminant function analysis was performed by W. J. S.

The data have been analyzed in three different but related ways.

The first is a measure-by-measure comparison which gives the nature of sex differences for each individual measurement. Visual graphs and standard statistical comparisons are used to report these. Second, a sorting procedure is followed whereby those individuals not separated by the ischium-pubis index method are tested by other methods. Twenty percent of our sample of 200 are not correctly sexed by the ischium-pubis index. When these 40 individuals are examined, other measurements are found which sex up to 95% of these. Last, but most important, a discriminant analysis is performed. With this statistical tool the aggregate values for sex discrimination of all measures taken together or in various combinations can be found. Everyone experienced in problems of sexing the skeleton knows that gross size, evidence of amounts of musculature and pelvic morphology are useful, but while these various indicators of sex are closely correlated they do have a measure of independence. Usually the more signs of sex that are evaluated the higher will be the likelihood of correct determination. It follows that each measure on the average has its own certain independent value and that it is the sum of these values that increases the likelihood. Discriminant analysis leads to the selection of a linear function of these measurements, which has maximum utility for assigning sex to individual specimens. The result is to give a simple formula for determining the sex of any single specimen, and a statement of the probability that sexing is correct for all the specimens in a series.

RESULTS

In figures 1-4, histogram graphs for femur length, femur head diameter, sternum width, epicondylar width of the humerus, clavicle length, humerus length, ischium length and pubis length are given. Figure 5 presents the graphs for ischium-pubis index.

The reader may visually evaluate the patterns of variability expressed in these figures. Of special interest is the fact that for pubis length the mean value for females is absolutely and significantly larger than the male value. While the ischium-pubis index was devised by Schultz ('30) it has been pointed out by others (Washburn, '48, '49, Hanna and Washburn, '53) that the pubis length relative to ischium is of great use in sexing human skeletons. If the measurement of the length of the pubis is divided by some measure which is highly correlated with gross body size, such as femur length, or more conveniently,

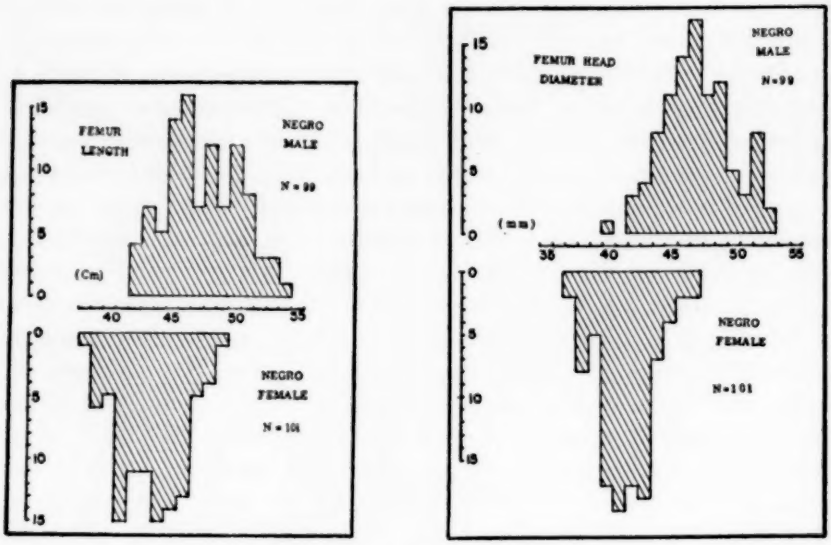

FIG. 1. THE DISTRIBUTION BY SEX OF MEASUREMENTS OF FEMUR LENGTH AND FEMUR HEAD DIAMETER.

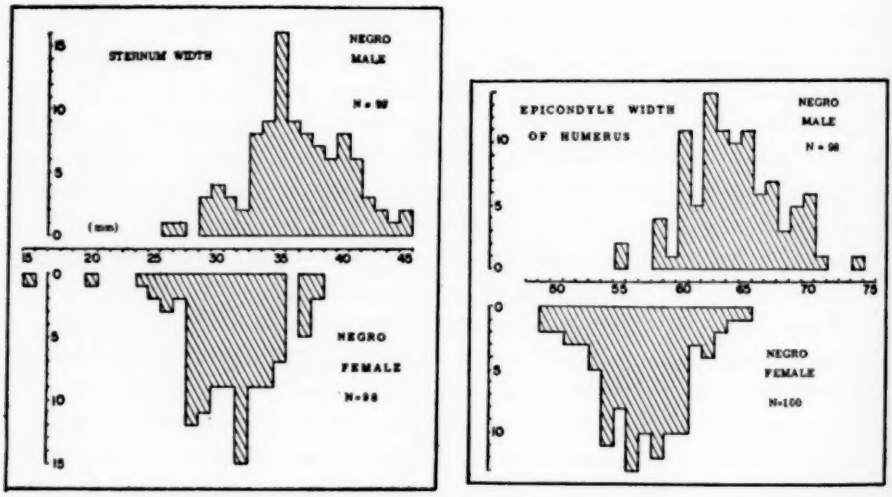

FIG. 2. THE DISTRIBUTION BY SEX OF MEASUREMENTS OF STERNUM WIDTH AND EPICONDYLAR WIDTH OF THE HUMERUS.

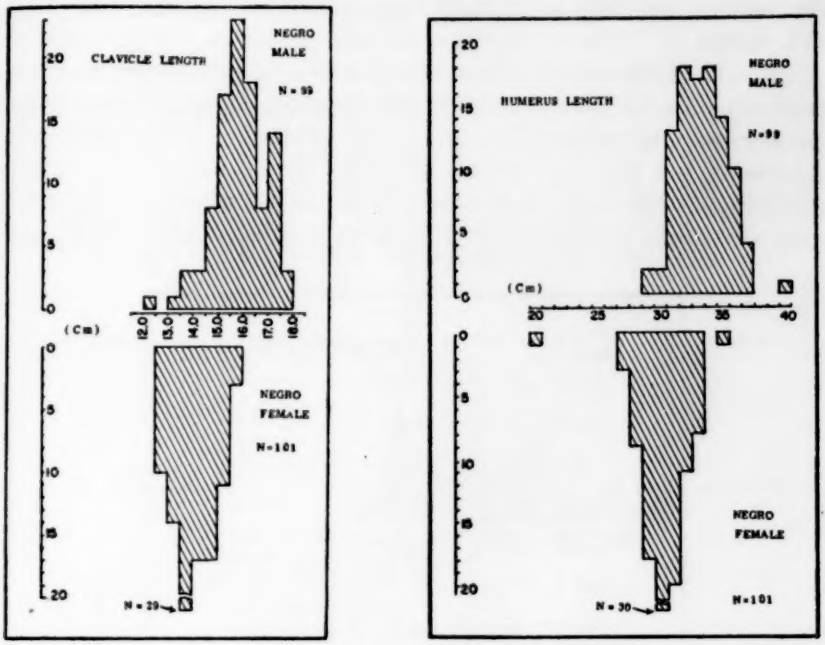

FIG. 3. THE DISTRIBUTION BY SEX OF MEASUREMENTS OF CLAVICLE LENGTH AND HUMERUS LENGTH.

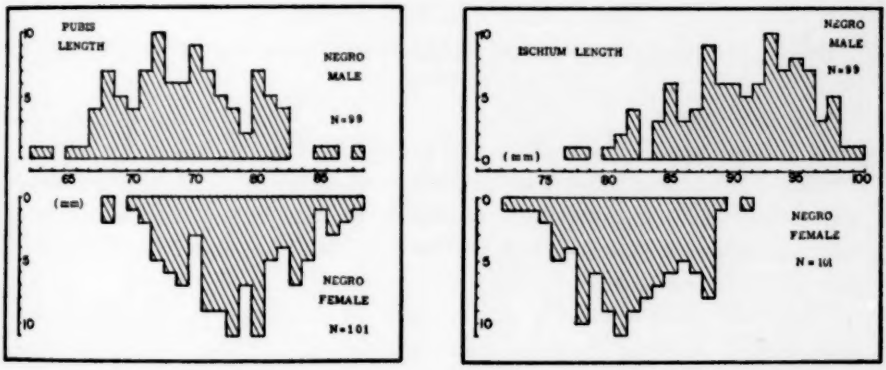

FIG. 4. THE DISTRIBUTION BY SEX OF MEASUREMENTS OF PUBIS LENGTH AND ISCHIUM LENGTH.

the ischium, an index with higher values for females than for males will result.

In table 1, the mean values, standard deviations, standard errors, male to female "*t*" values and the "differences of logs" for each of the measurements in the series are given. All of the "*t*" values are highly significant, but they are not the same for each measurement. While sternum width for instance, is relatively the most different (see figure 2) when the means are compared, it varies more widely (see table 1) and

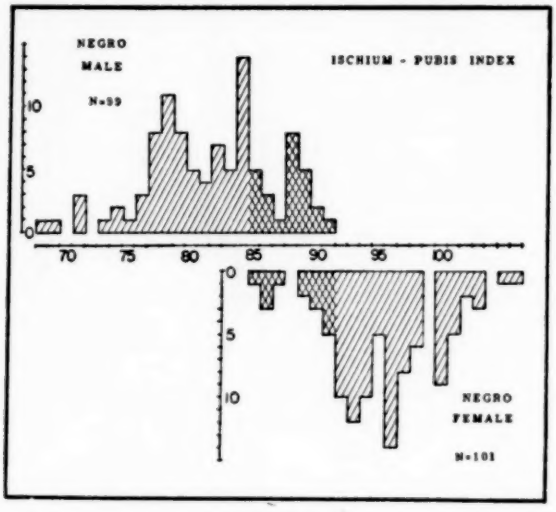

FIG. 5. THE DISTRIBUTION BY SEX OF MEASUREMENTS OF THE ISCHIUM-PUBIS INDEX.

thus is less significant because the standard error is relatively large. Pubis length shows a lower value than 6 of the other 7 measurements, but it is unique, as previously mentioned, in that it is the only one in which the female values exceed the male. Of all the measures used in this study femur head diameter is the best single discriminator.

Figure 6 gives a graphic arrangement of the mean values from the data and is designed to point out the levels of relative differences in overall form between the sexes. The method (Simpson, '41) expresses the difference between the logarithms of the means. One of the forms being compared (males) is arbitrarily given a value of zero. The difference between the logarithms of the mean value is plotted in

reference to this line. If the two forms are merely different in size but identical in proportions the lines will be parallel. If one is not merely the miniature of the other, however, the lines will not be parallel, but will be dissimilar to the degree that form differs. We see then

TABLE 1

Statistical values by sex for various measurements of the skeletons of American Negroes, and including results of tests of significance of difference between the sexes for each measurement and difference of log

	MEASUREMENT	*N*	SEX	MEAN (mm)	S.D.	S.E.$_m$	"*t*"	DIFFERENCE OF LOG
A	Femur length	98	♂	477.34	28.37	2.866	10.13	—.344
		100	♀	439.10	24.55	2.456		
B	Femur Head diam.	98	♂	47.17	2.75	.278	16.17	—.554
		100	♀	41.52	2.12	.212		
C	Humerus length	98	♂	338.98	18.55	1.874	12.51	—.45
		100	♀	305.89	18.66	1.866		
D	Epicondylar width of Humerus	98	♂	63.89	3.59	.363	14.50	—.51
		100	♀	56.76	3.32	.332		
E	Clavicle length	98	♂	158.24	10.06	1.158	13.90	—.52
		100	♀	140.28	7.99	.800		
F	Ischium length	98	♂	90.55	5.06	.511	13.52	—.44
		100	♀	81.80	4.00	.400		
G	Pubis length	98	♂	74.14	5.11	.516	6.39	+.23
		100	♀	78.21	4.34	.434		
	Sternum width	99	♂	36.15	3.91	.393	9.66	—.67
		98	♀	30.96	3.63	.367		

from figure 6 that females are not merely smaller than males, but each measure is relatively different. Consequently, some measurements are bound to be better than others in discriminating for sex.

Figures 1-6, while giving the picture of variation by sex for each of the 8 measurements, do not allow the reader to follow any individual and see his relative position in the scale of variability from graph to

graph. *Our problem is to sex the individual and not the measurement.* Short of publishing all the measurements for each individual, few methods will allow this sort of comparison to be made. But in view of the importance of following this point of attack, we give a partial analysis along this line. An examination of figures 1-5 shows that the

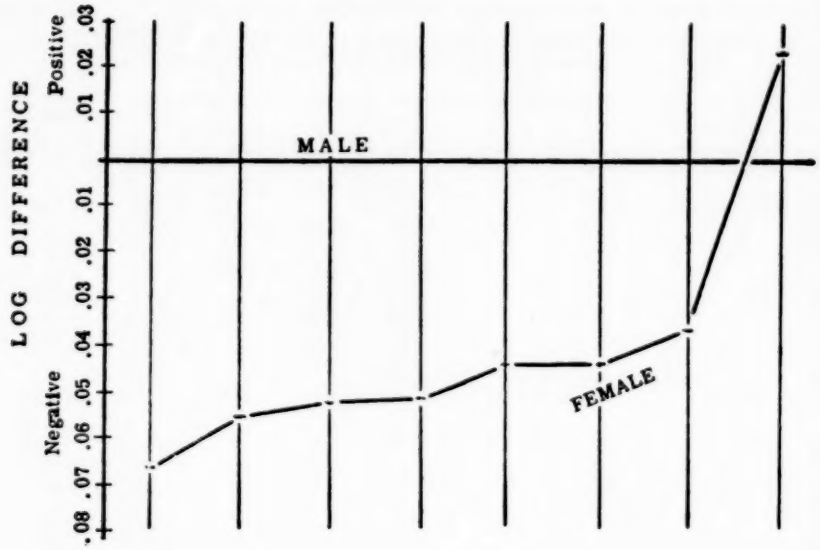

FIG. 6. COMPARISONS OF MALE-FEMALE MEAN VALUES FOR VARIOUS MEASUREMENTS USING THE DIFFERENCE OF LOG METHOD (SIMPSON, '41). The male values are located arbitrarily at 0 so the graph shows the relative difference of females compared to the males in terms of the difference between the logs of the means for each measurement.

ischium-pubis index by itself gives the best sorting. Forty of 200 individuals (20%) are not correctly sorted if we consider all with an index less than 85 to be male and all over 91 to be female (darkly shaded area, figure 5). This leaves 40 individuals not sexed by this single method. In figure 7 these 40 are placed according to their size for the 6 other measurements and sciatic notch opening. The ischium-pubis index is also shown. In addition, a particular symbol is assigned to those individuals who are not sexed by any one measure so that they may be followed to see how they fare in each array. Where the same symbol appears repeatedly it is apparent that for several measure-

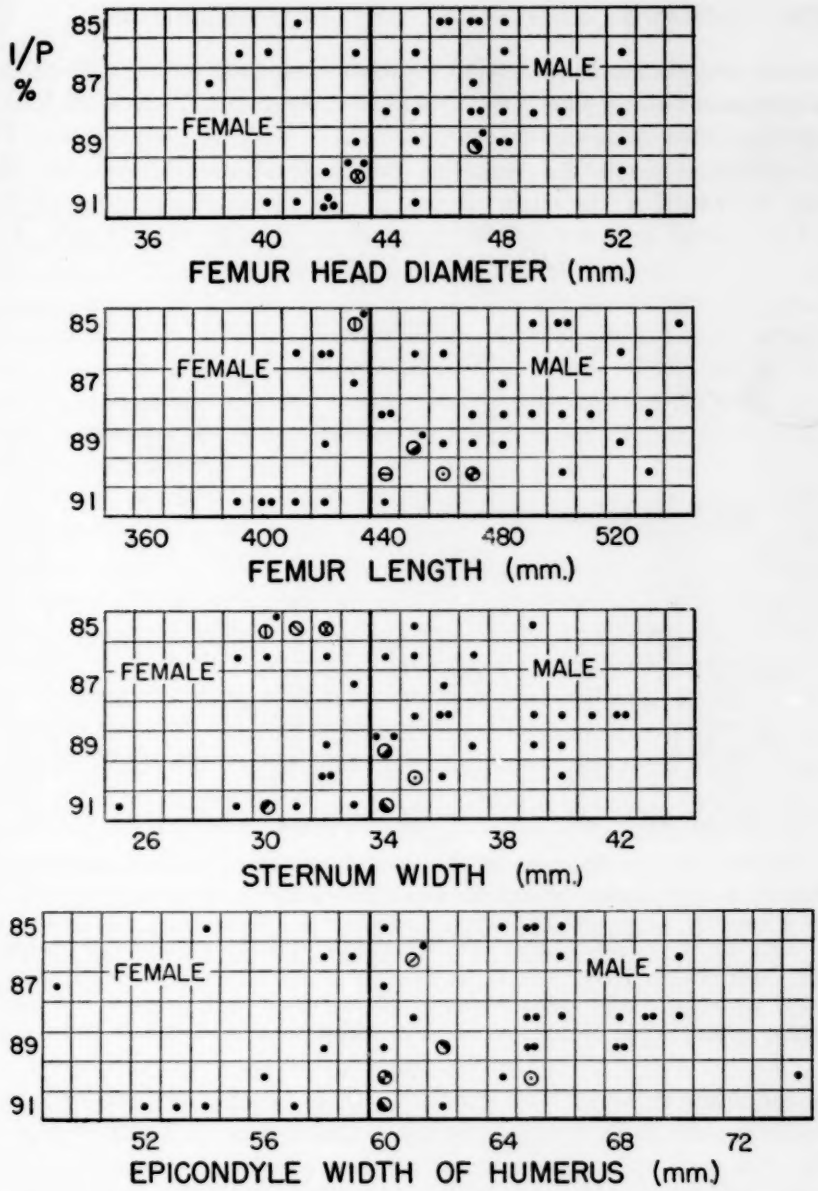

FIG. 7a. THE DISTRIBUTION OF 40 INDIVIDUALS (25 FEMALES AND 15 MALES SHOWN IN THE DARKLY SHADED AREA OF FIG. 5), WHOSE *I/P* INDEX IS BETWEEN 85.0 AND 92.0, ACCORDING TO OTHER MEASUREMENTS. Each case is indicated with a solid dot if properly located with reference to a heavy vertical line dividing the sexes in each array. The larger symbols denote cases which are not within their sex range. They are individually assigned except that ⊗ denotes all cases which appear out of their proper range only once.

ments that person is out of the expected size range for his sex. This arrangement shows how some individuals who are difficult to sex vary for the different measurements.[1]

The best single discriminating measurement next to the ischium-

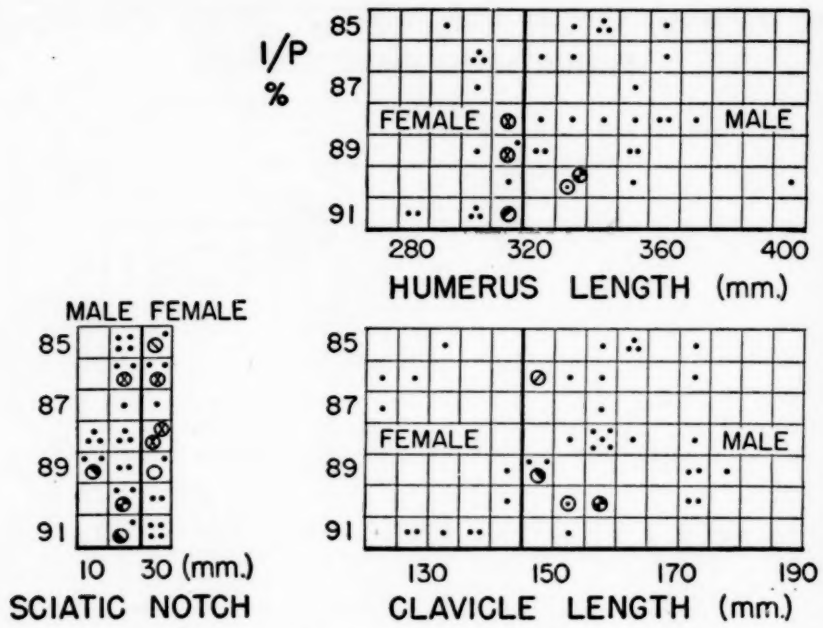

FIG. 7b. THE DISTRIBUTION OF 40 INDIVIDUALS (25 FEMALES AND 15 MALES SHOWN IN THE DARKLY SHADED AREA OF FIG. 5), WHOSE *I/P* INDEX *is* BETWEEN 85.0 AND 92.0, ACCORDING TO OTHER MEASUREMENTS. Each case is indicated with a solid dot if properly located with reference to a heavy vertical line dividing the sexes in each array. The larger symbols denote cases which are not within their sex range. They are individually assigned except that ⊗ denotes all cases which appear out of their proper range only once.

pubis index is femur head diameter. Only two persons not sexed by ischium-pubis index remain unsorted if a limit value of 44 mm diameter is used, with all under being female and all above being male. The method as presented here cannot be carried one step further, however.

[1] The symbol ⊗ is used to signify the dimensions of all those individuals who are out of the expected range for their sex in only one of the measurements shown. The other symbols are individually assigned to cases which more than once are out of the range for their sex.

Even though the symbols of the two individuals not sorted by femur head diameter fail to appear in the femur length or humerus length arrays, it does not follow that these two deviant individuals could be picked out of the sample if sex were unknown. Despite this, it is quite accurate to conclude that *in this sample* where sex is known we can sex with 99% accuracy if we first sort for ischium-pubis index, calling all individuals male if under 85.0 mm and all female who are 92.0 or over, and next sort those remaining, calling all under 44 mm for femur head female and all others male. The resulting male-female groups have only one individual per 100 incorrectly sexed. While we can be confident that this formula would probably give high accuracy for any other American Negro population or possibly any Negro population, we do not know the exact level (Thieme, in press).

A discriminant function analysis has been carried out using these data in order to determine the power that all the measures taken together have for discrimination. Discriminant analysis for problems of sexing skeletal material has been used previously (Mukherjee, Rao and Trevor, '55, Pons, '55). Rao applied the method to pre-historic skeleton material of unknown sex and Pons used several measurements on each of two bones, femur and sternum, but did not combine them for both bones to find their cumulative discriminating value. Using four measurements on the femur he found the sexing error to be about 5%, with femur head diameter and bicondylar width contributing most. These two papers, while demonstrating the value of discriminant analysis for sexing, do not show to what degree the method may be useful in multiple measurement analysis from several bones on material of known sex. In addition, while this is a valuable statistical method, it is not well known in physical anthropology. Because of the importance of this method, the following section gives the method involved in discriminant analysis in some detail and is arranged in step-by-step order so that it may have the highest practical value.

DISCRIMINANT ANALYSIS

The basic theory of discriminant analysis was developed by Fisher ('36, '38, '40) in a series of three classic papers. A number of other individuals, notably Mahalanobis ('36) and Rao ('52), have made significant contributions to the theory of parametric discriminatory topology, and recently, Fix and Hodges ('51) have investigated non-parametric

discriminant functions. We shall indicate only briefly the argument underlying parametric discriminant functions and then present in detail the computational procedure for constructing a discriminant function of the type to be employed on the present data. A complete presentation will be found in Rao ('52).

Generally the use of a discriminant function arises in the following context: We are given samples of size N_1 and N_2 drawn from two populations each characterized by p variates, and are required to find a linear function, say

$$X = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_p x_p$$

which will maximize the ratio of the between population to the within population variance of X for the two samples. That is to say, we are required to estimate $\lambda_1, \cdots, \lambda_p$ such that the ratio

$$\frac{[N_1 + N_2 - p - 1]}{p} \frac{\sum_i N_i (\bar{X}_i - \bar{X})^2}{\sum_{ij} (X_{ij} - \bar{X}_i)^2}$$

is a maximum, where $\bar{X}_i$ = the mean of the linear function in the i^{th} sample (here $i = 1, 2$), $\bar{X}$ = the grand mean of the linear function, and X_{ij} = the value of the linear function for the j^{th} individual in the i^{th} sample, and where $j = 1, \cdots, N_i$. The ultimate purpose of this linear function is, in most instances, to form a basis for assigning (classifying) unknown observations to one or the other of the two populations.

To give meaning to this ratio in experimental situations, we make two assumptions, namely, that

(1) each sample was drawn from a p-variate normal population, and

(2) the populations involved have homogeneous, *i. e.*, equal, variance-covariance matrices.

The construction of the discriminant function proceeds as follows (the reader can easily extend the computations to those cases where the variables are more numerous than three):

Step 1. The computation of means, differences between means, and the sum of products of deviations matrix: Available to us are the sums, sums of squares, and sums of cross products for the measurements femur head diameter (B), ischium length (F), and pubis length (G)

for samples of 98 Negro males and 100 Negro females. These values are

Males

$$\sum_{\sigma} B_{\sigma} = 4623 \qquad \sum_{\sigma} B_{\sigma}^2 = 218821 \qquad \sum_{\sigma} B_{\sigma} F_{\sigma} = 419410$$
$$\sum_{\sigma} F_{\sigma} = 8874 \qquad \sum_{\sigma} F_{\sigma}^2 = 806062 \qquad \sum_{\sigma} B_{\sigma} G_{\sigma} = 343423$$
$$\sum_{\sigma} G_{\sigma} = 7266 \qquad \sum_{\sigma} G_{\sigma}^2 = 541282 \qquad \sum_{\sigma} F_{\sigma} G_{\sigma} = 659236$$

Females

$$\sum_{\varphi} B_{\varphi} = 4152 \qquad \sum_{\varphi} B_{\varphi}^2 = 172842 \qquad \sum_{\varphi} B_{\varphi} F_{\varphi} = 340163$$
$$\sum_{\varphi} F_{\varphi} = 8180 \qquad \sum_{\varphi} F_{\varphi}^2 = 670726 \qquad \sum_{\varphi} B_{\varphi} G_{\varphi} = 325158$$
$$\sum_{\varphi} G_{\varphi} = 7821 \qquad \sum_{\varphi} G_{\varphi}^2 = 613561 \qquad \sum_{\varphi} F_{\varphi} G_{\varphi} = 640901$$

where $\sum_{\sigma}$ and $\sum_{\varphi}$ indicate that summation is over all males and all females, respectively. We now compute the following quantities for the males:

$$\bar{B}_{\sigma} = \frac{\sum B_{\sigma}}{N_{\sigma}} = 47.173469; \qquad S_{B\sigma} = \frac{N_{\sigma}\sum B_{\sigma}^2 - (\sum B_{\sigma})^2}{N_{\sigma}} = 738.06;$$
$$S_{BF\sigma} = \frac{N_{\sigma}\sum B_{\sigma}F_{\sigma} - (\sum B_{\sigma})(\sum F_{\sigma})}{N_{\sigma}} = 792.63$$

$$\bar{F}_{\sigma} = \frac{\sum F_{\sigma}}{N_{\sigma}} = 90.551020; \qquad S_{F\sigma} = \frac{N_{\sigma}\sum F_{\sigma}^2 - (\sum F_{\sigma})^2}{N_{\sigma}} = 2512.24;$$
$$S_{BG\sigma} = \frac{N_{\sigma}\sum B_{\sigma}G_{\sigma} - (\sum B_{\sigma})(\sum G_{\sigma})}{N_{\sigma}} = 660.57$$

$$\bar{G}_{\sigma} = \frac{\sum G_{\sigma}}{N_{\sigma}} = 74.142857; \qquad S_{G\sigma} = \frac{N_{\sigma}\sum G_{\sigma}^2 - (\sum G_{\sigma})^2}{N_{\sigma}} = 2560.00;$$
$$S_{FG\sigma} = \frac{N_{\sigma}\sum F_{\sigma}G_{\sigma} - (\sum F_{\sigma})(\sum G_{\sigma})}{N_{\sigma}} = 1294.29.$$

Similar quantities are then computed for the females. From these two arrays we form (1) the differences in the means

$$d_B = \bar{B}_{\sigma} - \bar{B}_{\varphi} = 47.173469 - 41.520000 = 5.653469$$
$$d_F = \bar{F}_{\sigma} - \bar{F}_{\varphi} = 90.551020 - 81.800000 = 8.751020$$
$$d_G = \bar{G}_{\sigma} - \bar{G}_{\varphi} = 74.142857 - 78.210000 = -4.067143$$

and (2) the elements of the sum of the products of deviations matrix

$$S_{BB} = S_{B\sigma} + S_{B\varphi} = 738.06 + 450.96 = 1189.02$$
$$S_{FF} = S_{F\sigma} + S_{F\varphi} = 2512.24 + 1602.00 = 4114.24$$
$$S_{GG} = S_{G\sigma} + S_{G\varphi} = 2560.00 + 1880.59 = 4440.59$$
$$S_{BF} = S_{BF\sigma} + S_{BF\varphi} = 792.63 + 529.40 = 1322.03$$
$$S_{BG} = S_{BG\sigma} + S_{BG\varphi} = 660.57 + 430.08 = 1090.65$$
$$S_{FG} = S_{FG\sigma} + S_{FG\varphi} = 1294.29 + 1143.20 = 2437.49.$$

Step 2. The sum of products of deviations matrix and the computation of the λ's. We now construct the symmetric matrix whose elements are

S_{BB}	S_{BF}	S_{BG}	d_B	1189.02	1322.03	1090.65	5.653469
	S_{FF}	S_{FG}	d_F		4114.24	2437.65	8.751020
		S_{GG}	d_G			4440.59	—4.067143.

To this matrix we adjoin a unit matrix, and the extended matrix so formed appears as follows:

							Sum
1189.02	1322.03	1090.65	5.653469	1			3608.353469
(1322.03)	4114.24	2437.65	8.751020	0	1		7883.671020
(1090.65)	(2437.65)	4440.59	—4.067143	0	0	1	7965.822857.

The column headed "Sum" is obtained by summing across all of the elements in the row.

To obtain the estimates of the λ's we are required to reduce this extended matrix. A convenient method of reduction is the "square root method" (for a detailed presentation of this and other methods of reduction and a description of the errors in these methods the reader is referred to Dwyer, '51). To reduce the matrix by this method we proceed as follows: We may ignore the terms in parentheses in the above matrix, and compute the matrix whose elements are as indicated below:

s_{11}	s_{12}	s_{13}	d_1	t_{11}	0	0
	$s_{22\cdot1}$	$s_{23\cdot1}$	d_2	t_{21}	t_{22}	0
		$s_{33\cdot12}$	d_3	t_{31}	t_{32}	t_{33}

where the s elements are obtained by operating on the S's, the d elements

on d_B, etc., and the t elements from the unit matrix. The s, d, and t elements are calculated as follows:

$$s_{11} = \sqrt{S_{BB}} = \sqrt{1189.02} = 34.482169$$

$$s_{12} = \frac{S_{BF}}{s_{11}} = \frac{1322.03}{34.482169} = 38.339526$$

$$s_{13} = \frac{S_{BG}}{s_{11}} = \frac{1090.65}{34.482169} = 31.629391$$

$$d_1 = \frac{d_B}{s_{11}} = \frac{5.653469}{34.482169} = 0.163953$$

$$t_{11} = \frac{1}{s_{11}} = \frac{1}{34.482169} = 0.029000$$

$$s_{22\cdot1} = \sqrt{S_{FF} - (s_{12})^2} = \sqrt{4114.24 - (38.339526)^2} = 51.422959$$

$$s_{23\cdot1} = \frac{S_{FG} - (s_{12})(s_{13})}{s_{22\cdot1}} = \frac{2437.65 - (38.339526)(31.629391)}{51.422959} = 23.821930$$

$$d_2 = \frac{d_F - (s_{12})(d_1)}{s_{22\cdot1}} = \frac{8.751020 - (38.339526)(0.163953)}{51.422959} = 0.047939$$

$$t_{21} = \frac{0 - (s_{12})(t_{11})}{s_{22\cdot1}} = \frac{0 - (38.339526)(0.029000)}{51.422959} = -0.021622$$

$$t_{22} = \frac{1}{s_{22\cdot1}} = \frac{1}{51.422959} = 0.019447$$

and

$$s_{33\cdot12} = \sqrt{S_{GG} - (s_{13})^2 - (s_{23})^2}$$

$$d_3 = \frac{d_G - (s_{13})(d_1) - (s_{23})(d_2)}{s_{33\cdot12}}$$

$$t_{31} = \frac{0 - (s_{13})(t_{11}) - (s_{23})(t_{21})}{s_{33\cdot12}}$$

$$t_{32} = \frac{0 - (s_{23\cdot1})(t_{22})}{s_{33\cdot12}}$$

$$t_{33} = \frac{1}{s_{33\cdot12}}.$$

It should be noted that the computations for the second and third (or subsequent) rows are analogous. In the present example, we obtain the following values:

							Check
34.482169	38.339526	31.629391	0.163953	0.029000	0	0	104.644040 104.644039
	51.422959	23.821930	0.047939	—0.021622	0.019477	0	75.290652 75.290653
		53.597456	—0.193943	—0.007504	—0.008643	0.018658	53.406023 53.406024.

The rightmost column in the above matrix affords a check of the computations. It will be noticed that each row has two entries in this column. The first entry is computed from the sum in the original extended matrix by the same method used to generate the other entries in the row, thus

$$\text{first row} = \frac{3608.353469}{34.482169} = 104.644040$$

$$\text{second row} = \frac{7883.671020 - (104.644040)(38.339526)}{51.422959} = 75.290652$$

$$\text{third row} = \frac{7965.822857 - (104.644040)(31.629391) - (75.290652)(23.821930)}{53.597456} = 53.406023.$$

The second entry is obtained by summing the elements across the row, for each row, in the reduced matrix. These entries should be equal, or nearly so, if the computations have been correctly performed.

Now to compute the estimates of the λ's and their variances we employ terms obtained from the reduced matrix. It can be shown that

$$\lambda_3 = \sum_i d_i t_{i3} = (-0.193943)(0.018658) = -0.003619$$

$$\lambda_2 = \sum_i d_i t_{i2} = (-0.193943)(-0.008643) + (0.047939)(0.019447) = 0.002609$$

$$\lambda_1 = \sum_i d_i t_{i1} = (-0.193943)(-0.007504) + (0.047939)(-0.021622) + (0.163953)(0.029000) = 0.005173.$$

Step 3. The significance of the discriminant function and the errors associated with the λ's. Once the values of the λ's associated with the linear compound are computed we must determinine whether the discriminant function is effective. Otherwise stated, we are concerned to determine whether significant discrimination is achieved. The significance of the discriminant function can be ascertained by an analysis of variance. We form the analysis of variance table

SOURCE	SUM OF SQUARES	DF	M.S.	F
Between sexes	$\sum_i N_i(\bar{X}_i - \bar{X})^2 = \frac{N_1 N_2}{N_1 + N_2}[\sum_i \lambda_i d_i]^2$	p	A	A/B
Within sexes	$\sum_{ij}(X_{ij} - \bar{X}_i)^2 = \sum_{ij}\lambda_i\lambda_j S_{ij}$	$N_1 + N_2 - p - 1$	B	
Total	$\sum_{ij}(X_{ij} - \bar{X})^2$	$N_1 + N_2 - 1$		

In the illustrative example, we find

SOURCE	SUM OF SQUARES	DF	M.S.	F
Between sexes	0.220776	3	0.073592	213.930
Within sexes	0.066798	194	0.000344	
Total	0.287574	197		

The F is significant for 3 and 194 degrees of freedom, indicating that significant discrimination is achieved. Were F not significant, the discriminant function would be ineffective in classifying sex.

A further measure of the effectiveness of the discriminant function can be gleaned from a comparison of the individual λ's with their standard errors (see Kendall, '48). At this juncture, it should be pointed out that the approach to be described is not without difficulty since the coefficients, the λ's, of the discriminant function are not unique in the sense that they are estimates of specific population parameters, only their ratios are unique.

The standard errors of the coefficients we compute from the mean square within samples and elements obtained from the reduced matrix obtained in step 2. It can be shown that

$$\sigma_{\lambda_3}^2 = 0.000344 \sum_{i=1}^{3}(t_{3i})^2 = 0.000344(0.018658)^2 = 0.000000119754$$

$$\sigma_{\lambda_2}^2 = 0.000344 \sum_{i=1}^{3}(t_{2i})^2 = 0.000344[(-0.008643)^2 + (0.019447)^2]$$
$$= 0.000000177593$$

$$\sigma_{\lambda_1}^2 = 0.000344 \sum_{i=1}^{3}(t_{1i})^2 = 0.000344[(-0.007504)^2 + (-0.021622)^2 + (0.029000)^2]$$
$$= 0.000000469498$$

and hence

$$\sigma_{\lambda_1} = 0.00068520$$

$$\sigma_{\lambda_2} = 0.00039471$$

$$\sigma_{\lambda_3} = 0.00034605.$$

Since each of the λ's exceeds twice its standard error, we may assert that each of the measurements contributes to the effectiveness of the discriminant function. If one or more of the λ's did not exceed twice its standard error, this would imply that the measurements associated with these λ's did not contribute significantly to the discriminant function. These measurements could then be discarded, and a new discriminant function computed. In practice, this would amount to recomputing the entries in the rows following the row corresponding to the first discarded measurement. If the last row (or rows) is omitted, no further computation is necessary. A corollary of this would lead one to place in the last row that measurement which on an *a priori* basis would seem to contribute the least to the discriminant function.

Step 4. The discriminating value: To assign individuals to one of the two populations sampled we need a standard of judgment. This standard we shall term the discriminating value, and shall assign to one population values which exceed this standard and to the other population values less than this standard. We compute the discriminating value as follows: We determine, first, the means of the linear function associated with the two populations, thus

$$\begin{aligned}\bar{X}_{\sigma} &= \lambda_1\bar{B}_{\sigma} + \lambda_2\bar{F}_{\sigma} + \lambda_3\bar{G}_{\sigma} \\ &= (0.005173)(47.173469) + (0.002609)(90.551020) \\ &\quad + (-0.003619)(74.142857) \\ &= 0.211953\end{aligned}$$

$$\bar{X}_{\varphi} = \lambda_1\bar{B}_{\varphi} + \lambda_2\bar{F}_{\varphi} + \lambda_3\bar{G}_{\varphi} = 0.145157.$$

To obtain the discriminant value, $\bar{X}$, we merely solve

$$\bar{X} = \frac{\bar{X}_{\sigma} + \bar{X}_{\varphi}}{2} = 0.178555.$$

Accordingly, we shall assign to the male population those individuals

where the discriminant function exceeds 0.178555, and to the female population those individuals where the discriminant function is less than 0.178555. A discussion of the rationale underlying the choice of the discriminating value will be found in Rao ('52, p. 296).

In scoring individuals, it is often convenient to convert the λ's to values more easily remembered. The actual values are relatively unimportant since discrimination rests on the proportional relationships. We may, then, divide λ_1, λ_2, and λ_3 by a constant factor, say λ_2, which would make $\lambda'_1 = \lambda_1/\lambda_2$, $\lambda'_2 = 1$, and $\lambda'_3 = \lambda_3/\lambda_2$, and in the present example, $\lambda'_1 = 1.9828$, $\lambda'_2 = 1.0000$, and $\lambda'_3 = -1.3871$. This device, of course, leads to a change of scale in the discriminant function with the new discriminating value being 68.42.

Step 5. The classificatory error: We have indicated that the purpose of a discriminant function is to facilitate assigning to one of two or more populations an observation known to have been drawn from one, unspecified member of this group of populations. If the populations involved are uniquely different, obviously no discriminant function is necessary. More frequently, however, the populations are not unique and hence there arises the possibility of misclassification of an observation. One of the advantages of the discriminant function is the opportunity which it affords to appraise the classification error in an objective fashion, an advantage not enjoyed by subjective or unquantified experimental standards.

To compute the classification error associated with a discriminant function we proceed as follows, again using the illustrative material: We have indicated in the previous section that we shall assign an observation to the male population if the value of the discriminant associated with the observation exceeds 0.178555, to the female population if the value of the discriminant function is less than 0.178555, and in the unlikely event that the discriminant value observed was precisely equal to 0.178555 we would assign the observation to the male or female group by tossing a coin (if λ'_1, λ'_2, and λ'_3 are used then 0.178555 must be replaced by 68.42). Now we should expect, by chance, to obtain some male values which were smaller than 0.178555 and would, therefore, be erroneously assigned to the female population, and conversely, we should expect to obtain some female values which were larger than 0.178555 and which would be erroneously assigned to the male population. Our task is to determine how frequently these two

events would occur. As a consequence of the assumption that each sample is drawn from a p-variate normal population, the observed discriminant values, for a given population, will be normally distributed around the mean value for that population, hence Negro male values will be normally distributed about the mean 0.211953, and female values around 0.145157. The error frequency associated with females, then, will correspond to the area to the right of 0.178555 of a normally distributed variable with mean 0.145157, and variance equal to 0.00034432 (mean square within sexes).

Now the normal deviate with zero mean and unit variance is merely

$$\frac{0.178555 - 0.145157}{0.00034432} = \frac{0.033398}{0.018556} = 1.800.$$

From a table of the normal deviate, we note that 3.6% of the area lies to the right of this value. We should, then, misclassify females with a frequency of roughly four in one hundred. The frequency with which males would be misclassified would be the same since the discriminating value is so chosen as to equalize the classificatory errors.

Suppose, now, that an unknown specimen was drawn from the 198 individuals available to us, and we ask how frequently would we misclassify this "unknown" individual. This frequency would merely be the sum of the products of (1) the probability of drawing a male and the probability of misclassifying him, and (2) the probability of drawing a female and the probability of misclassifying her. Since the frequency of males and females is approximately 1:1 in this group, this sum is merely 3.6%.

DISCRIMINANT ANALYSIS RESULTS

In carrying out discriminant analysis it is necessary to get sums of the cross products for each individual pair of the measures used. Inasmuch as two of our series, one male and one female, could not be measured for the epicondylar width of the humerus, the calculations for the statistics for each measurement and for the discriminant function are based upon a reduced sample (98 male and 100 female). The figures 1-7 include the whole series. The measurements of these two individuals were not in the 'problem' range for the ischium-pubis index and their removal from the sample has little effect. (Sample used is shown in Table 1).

Table 2 gives the correlation coefficients between all of the 7 measurements used. (Sternum width was not used as this would have reduced the female sample even further and because it is not a very good discriminator, as can be seen in figure 2).

TABLE 2

Correlation coefficients between all pairs of measurements used in the discriminant analysis

MEASUREMENT	A	B	F	G	C	D	E
A	1.0000						
B	.4627	1.0000					
F	.6477	.5983	1.0000				
G	.6011	.4753	.5708	1.0000			
C	.8108	.4280	.5902	.5086	1.0000		
D	.4179	.5356	.5083	.4370	.4212	1.0000	
E	.5629	.2625	.4295	.4475	.5257	.2709	1.0000

where A = femur length
B = femur head diameter
F = ischium length
G = pubis length
C = humerus length
D = epicondylar width of humerus
E = clavicle length

Table 3 gives the coefficients of variation (proportion of variation in one variable explained by variation in another). In table 4 some results (3, 4, 6 and 7 measure combinations) of the discriminant analysis are given.

In figure 8 the data from this series are used to test how one set of values developed in the discriminant analysis apply to this particular series. The discriminant function for 4 measurements is calculated for each individual and the values plotted on a scale. Figure 8 shows that the totals for three males are less than 665, the limit value, and one female more. In other words, 4 individuals, or 2% of the series, are not accurately sexed. The probability of wrong classification of

3.1%, as calculated in the discriminant analysis, was in fact slightly bettered.

In the 6 measurement analysis, only two skeletons,[2] one female and one male, are not properly sexed when the values are applied to this series. The same two individuals are again not sexed in the 7 measurement series.

TABLE 3

Coefficients of variation (proportion of variation in one variable explained by variation in another) between all pairs of measurements used in the discriminant analysis

MEASUREMENT	A	B	F	G	C	D	E
A	1.0000						
B	.2141	1.0000					
F	.4195	.3580	1.0000				
G	.3613	.2259	.3258	1.0000			
C	.6574	.1832	.3483	.2587	1.0000		
D	.1746	.2869	.2584	.1910	.1774	1.0000	
E	.3169	.0689	.1845	.2002	.2764	.0734	1.0000

where A = femur length
B = femur head diameter
F = ischium length
G = pubis length
C = humerus length
D = epicondylar width of humerus
E = clavicle length

In practice the values given in table 4 are quite simple to use. Three, 4, 6, or 7 measurements are taken, the appropriate multipliers are used as shown in the table and the totals of these products are obtained. If the total is under the limit value, the skeleton is female; if over, it is male.

[2] The males with a femur head diameter of only 40 mm would have been correctly sexed if it had instead been 42 mm. The female, with a femur head of 47 mm would have been correctly sorted if it had been 45 mm. The ischium-pubis index for the male is 85.2 and the female 89.0, both in the "overlap area." It is in the unusual diameters of the femur head that these two skeletons are especially deviant.

TABLE 4

The results of the discriminant analysis for combinations of 3, 4, 6 and 7 measurements as given in multipliers to be used, the limit values, probabilities of misclassification and expected ranges

	7 × 7	6 × 6	4 × 4	3 × 3
A. Femur length	.07	1.00	1.00	
B. Femur head diameter	58.14	31.40	16.53	1.98
F. Ischium lenth	16.25	11.12	6.10	1.00
G. Pubis length	—63.64	—34.47	—13.80	—1.39
C. Humerus length	2.68	2.45		
D. Epicondylar width of humerus	27.68	16.24		
E. Clavicle length	16.09			

Discriminating values

	VALUE SEPARATING THE SEXES (♀ ARE LESS THAN THE VALUE)	PROBABILITY OF ERROR IN CLASSIFICATION PER CENT	RANGE WITHIN WHICH 99% WILL FALL
7 × 7	4099	1.5	2664-5533
6 × 6	1953	2.5	1205-2701
4 × 4	665	3.1	367-963
3 × 3	68	3.6	50-86

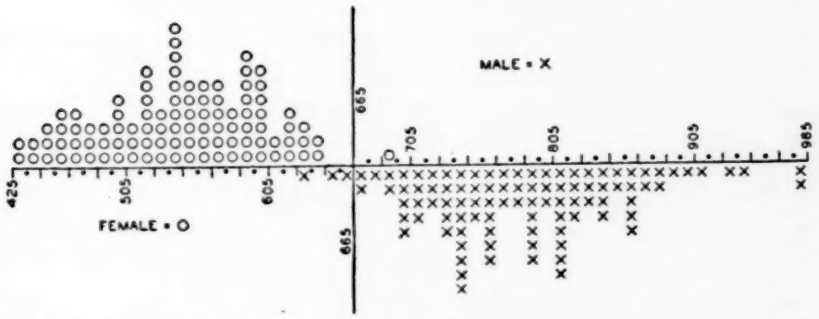

FIG. 8. THE RESULTS OF A TRIAL OF DISCRIMINANT ANALYSIS VALUES (4 × 4) APPLIED TO 198 AMERICAN NEGRO BONES OF KNOWN SEX.

For example, an actual case which cannot be sexed by ischium-pubis index or femur head size has the following measurements:

1. femur length	429
2. femur head diameter	44
3. ischium length	84
4. pubis length	73

The multipliers from table 4 give:

+ 1.00 × 429 =	429.00
+16.53 × 44 =	727.32
+ 6.10 × 84 =	512.40
−13.80 × 73 =	−1007.40
Total =	661.32

As the limiting value is 665 and all under it are female, this skeleton is classified correctly as female.

RIGHT-LEFT ASYMMETRY BY SEX

In table 5 the results are given by right-left comparisons for the femur, humerus and clavicle by sex.

TABLE 5

Dominance in length in right-left comparisons by sex in percentage for three bones from each of 200 American Negroes

MEASUREMENT	MALE			FEMALE		
	Right	Left	Equal	Right	Left	Equal
Femur	37.4	44.4	18.2	35.2	50.5	14.3
Humerus	44.0	42.0	14.0	71.4	14.3	14.3
Clavicle	23.0	65.0	12.0	38.1	52.4	9.5

There is a striking dominance for the right side in the humerus for females and for left dominance in male clavicles. While the pattern for both sexes is apparently right-armed and left-legged, the simple functional explanation that this implies does not hold for the clavicle. Trotter and Gleser ('52) find this same dominance pattern with the exception that in their Negro males series the left humerus was slightly longer than the right, but the right arm dominance was borne out for

radius and ulna. In our series the right is only slightly more frequent than the left.

While of no obvious value for sexing skeletons, these right-left differences are not the same for each sex and should be explained by something other than mere functional handedness dominance. The absolute size of the mean differences between sides is small with the largest being 4.5 mm in the left male clavicle.

DISCUSSION

The extent of sexual variability for several measurements is given in figures 1-6. Methods of sorting individual skeletons based on this variability pattern have also been presented. These methods, while highly efficient for this series of adult American Negroes, may have different value for sexing other series, or individuals from other populations. What degree of general applicability do these methods have? The answer to this depends first, upon the degree that gross size varies *between* populations for each sex, and second, on the degree that the sexual dimorphism in the skeleton is the same *within* these populations. The answer to the first question is obvious. Variation in size is marked between races, and consequently the absolute size of the limiting values separating the sexes must be revised for each race according to their relative place on a scale of racial variation in mean size for each measurement. For instance, while 44 mm may be the point in the scale of size of the head of the femur which best separates Negro males and females, it would undoubtedly be a smaller value for some other groups and larger for still others.

The second question dealing with the degree of regularity in the expression of relative sex differences within the varieties of mankind is not so easily answered. The fundamental pattern for sexual dimorphism is basically a species characteristic but we are concerned here with the degree that a racial variation exists within the species. We have no precise answer to this question. Schultz ('37) states, ". . . the various human races differ but very little in the degree with which the males surpass the females in the length of the long bones." He shows that the relative length of the female to male humerus or femur is between 91% and 93% for White, Negro, Eskimo, and American Indian series with an average of 91.84%. In Portuguese (Pons, '55) it is 89.5% and in our series it is 92.0. Hrdlička ('47) indicates little

or no sex differences in the ratio of femur length to stature for either the White or Negro races. Pearson (1899) in accumulating his basic data used for deriving curves for estimating stature from long bones was of the same opinion. In fact, the early workers dealing with the problem of estimating stature felt that possibly a single formula was equally efficient despite differences in race as well as sex. Recent studies (Dupertuis and Hadden, '51; Trotter and Gleser, '52) however, have shown that this is not the case and that each race probably has a slightly different formula describing the relationships. Trotter and Gleser go even further and find that there are differences for both race and sex. They state (p. 511), "It is evident . . . that it is necessary for the sake of obtaining the most accurate estimates of stature to have different equations for each of the two sexes and for each of the two (White and Negro) races." From their data we may conclude that the pattern of dimorphism, while not identical for these races, is only slightly different. In conclusion, there is a predominant similarity in the degree of sexual dimorphism for man with a minor but unknown amount of difference at the racial level.

Schultz ('30, '37, '49) has shown that in comparing the relative proportions by sex in the long bones of man and the apes there are appreciable differences among these primate forms. Also the ischium-pubis relations are quite different with a much greater overlap between the sexes in the ape index than in man. Washburn ('42) has shown that for three species of monkeys the index is different from man with no overlap occurring between the sexes. Schultz ('37) has also shown that the extent to which the sexes differ in size fluctuates widely in primates. He concludes that man stands in regard to the size of sex differences in between the macaque and chimpanzee on one side and the orang-utan and gorilla on the other. From these data describing differences in primate skeletons of known sex we may conclude that the expression of sexual dimorphism differs among primate forms. There is no basic primate pattern holding for all varieties, nor is there apparently a monkey pattern or an ape pattern, yet man seems to have a highly homogeneous pattern within the species.

In light of the previous discussion, we may draw the following conclusions. First, we can effectively operate at present on the assumption that the expression of sexual dimorphism is relatively similar in pattern for all varieties of man. This means that the shape of the bimodal curves describing sexually varying characters within popula-

tions will be approximately the same. Second, the location of these curves on a scale of absolute size will vary for different races. The limit values efficient for separating the sexes will consequently vary.

Consideration of these conclusions about the nature of the sexing problem suggests a practical application of our discriminant analysis beyond the American Negro population. An assumption of normal distribution is basic to the discriminant analysis. As our problem is to redefine the limit value for other populations, the following method is suggested, based upon the expectation that normal curves will describe the variability in any population for these measurements. If mean values for each combination of measurements as used in this study are known by sex for any other group, the multipliers can be used and an average value for each sex determined. The discriminating limit for any population will be one half the difference between these average values for each sex. Limit values so determined could be applied to skeletons of unknown sex with a high level of confidence.

Continuing to point out practical approaches to the problem of sexing skeletons, we might mention the results obtained when an analysis of discrimination for a redefined ischium-pubis index measurement is done. The ischium-pubis index as previously used has ill-defined landmarks and thus suffers in accuracy. The following is offered as an improvement.

If the pubic measurement were to be taken from the nearest acetabulum border to the superior point of the pubic symphysis and the ischium length to the far border of the acetabulum, the femur head size also would be included in the single index. Femur head diameter should be included because it is probably the best single measurement for sexing and would be included in this index because it is highly correlated with acetabulum size. To test roughly this idea, an index was made by taking pubis length minus half the femur head diameter dividing by ischium length plus half of femur head diameter. Using this method, a limit value of .500 correctly sexed all but 7 individuals in the series which is success at the 96.5% level. While this method is not precisely what was recommended, the limit value for this method will probably be close to .5 which would make it quite easy to use. A series of known sex pelves should be studied to determine what the real limit should be and how efficient it is.

Additional support of the idea that femur head should be used with the pubis in making a simple sexing index is the fact that in this

series a pubis-femur head index is more efficient than the ischium-pubis index. Using a value of 89.0 for the limiting value in ischium-pubis index, 13 individuals are not properly sexed, whereas using a limit value of 1.72 for pubis-femur head, only 11 are not properly sexed.

There are many "hallmarks" of sex that are proposed by various authors,[3] yet it seems that the efficiency achieved in sexing the skeleton by measurement or morphological form has about reached the upper limit. We notice in examining the nature of variability in the features recommended in the literature for use in measurement or observation that there are certain deviant individuals who are completely out of their expected range.

The variability for the "hallmarks" of sex in the skeleton is distributed along pathways separating only after an undifferentiated pre-puberty. Comparing the sexes, the males differ from pre-puberty most in changes in gross size of long bones, as can be seen from any growth curve. The females differ less in longitudinal growth but show special effects of endocrine mediated changes. Washburn ('48) in discussing this has shown that while the range of ischium-pubis index is the same for each sex before puberty, with puberty the female alone has appreciable growth in pubis length. Analysis of the patterns of variability by sex seems to indicate that the 'hard to sex' skeletons—the deviant individuals—are the males who did not achieve their expected size change and the females who exceeded expectation. The failure of individuals to conform to the expected endocrine patterns seems much less frequent. Without exception the standard error of the mean for male variability is greater than for females. The number of males to females in the "overlap" range for ischium-pubis index is 25:15. For Washburn's series ('48, '49) the ratio for American Negroes is 13:4, and for Bantu it is 19:3. From this it seems that the problem cases for sexing are the males who have interference with their normal longitudinal growth, and females who greatly exceed the average for their sex. Possibly a series of healthy individuals would give skeletons free of any problem cases. As most human skeletons of known sex come from dissecting room material we do not know to what extent the peculiarities of this sample source create difficulties in sexing.

[3] An example is given in a recent study of Sauter and Privat ('55). They summarize many previous methods as well as propose a new index. A more thorough summary of sexing procedures and a study of sex differences in the innominate bone suggesting new methods is given by Genoves ('55).

One clear conclusion to be drawn from this study is that sex in the skeleton can be determined with great accuracy. In an unknown series of American Negroes, using the 7×7 discriminant analysis values, over 98% accuracy is assured. The levels are lower for other methods but the results are still better than those obtained if sexing is done by using apparent total morphological configuration.

In view of the controversy over this point (Stewart '54), some remarks seem in order. In the first place it should be mentioned that the probability statements that are a property of a certain method apply only to series in which discrimination is done by that method. If a single specimen is to be sexed, the relative position of that individual on the total scale of variability for sex is what dictates the probability. By eye or by measure it may be readily known that there is little chance of incorrect sexing. For instance, in figure 8 the two males with values over 975 are *certainly* male. The point is not that one can be sure, but rather it is that the inexperienced person can be as sure as the expert. In the alternate cases where the signs of sex are confusing and the values are close to the limit value, the beginner or the expert can be equally certain of the doubtful nature of the case. It is most important that the beginner, and possibly the expert as well, should know when he is in trouble. The measurement method can tell him.

Stewart ('54) writes that attempts to sex by measurement are not better than the old methods which gave greatest weight to total morphological configuration. He interprets Washburn and Hanna ('53) to recommend that "measuring the pelvis should replace the traditional method in which the whole skeleton, or whatever part is available is inspected." (p. 389.) He concludes that only in special cases should it ever be necessary to measure and that this method be applied only when there is a certain basis for the need to use measurement. Stewart goes on to say. "It is a waste of time and effort to measure such specimens (the better differentiated ones) simply to verify what the eye so quickly discerned." It is undoubtedly true that measurement is more laborious but the crux of the matter is not this. The error lies in the fact that sexing by eye, although it uses one 'tool,' is not one method. Stewart's remarks are quite in order if he is commenting on the problems of sexing an isolated single specimen. On the other hand, for determining sex in a series it is obviously of value to use a single and explicit method. Different criteria should not be used for each specimen if it can be avoided. Otherwise there is no possibility of accurately

estimating error and an element of uncertainty of an indeterminate amount will color and dilute the security of any related conclusions.

The appeal of the measurement method is primarily derived from the degree that verifiable scientific methods are introduced to replace less satisfactory ones. There are cases where fragmentary evidence and unusual circumstances require the services of an expert who must almost intuitively arrive at conclusions, but this does not mean that such an approach is satisfactory when it is unnecessary. The introduction of precise measurement methods does not rule against the need for the expert but rather adds variety to his methods and demonstrable certainty to his conclusions.

LITERATURE CITED

DUPERTUIS, C. W., AND J. A. HADDEN, JR. 1951 On the reconstruction of stature from long bones. Am. J. Phys. Anthrop., n. s., *9*: 15-54.

DWYER, P. S. 1951 Linear Computations. Wiley and Sons, New York.

FISHER, R. A. 1936 The use of multiple measurements in taxonomic problems. Ann. Eugen., 7: 179-188.

——— 1938 The statistical utilization of multiple measurements. Ann. Eugen., *8*: 376-386.

——— 1940 The precision of discriminant functions. Ann. Eugen., *10*: 422-429.

FIX, EVELYN AND J. L. HODGES, JR. 1951 Discriminatory Analysis: Nonparametric Discrimination. USAF School of Aviation Medicine, Project No. 21-49-004, Reports No. 4 and 11.

GENOVES, SANTIAGO 1955 A study of sex differences in the innominate bone (os coxae). Thesis. Corpus Christi College, Cambridge.

HANNA, R. E. AND S. L. WASHBURN 1953 The determination of the sex of skeletons, as illustrated by a study of the Eskimo pelvis. Human Biol., *25*: 21-27.

HRDLIČKA, A. 1947 Practical Anthropometry. Third Edition, edited by T. D. Stewart. Wistar Institute, Philadelphia.

KENDALL, M. G. 1948 The Advanced Theory of Statistics, Volume 2. Charles Griffin and Co., London.

MAHALANOBIS, P. C. 1936 On the generalized distance in statistics. Proc. Nat. Inst. Sci. (India), *12*: 49-55.

MUKHERJEE, R., C. R. RAO AND J. C. TREVOR 1955 The Ancient Inhabitants of Jebel Moya. Cambridge University Press, Cambridge.

PONS, JOSÉ 1955 The sexual diagnosis of isolated bones of the skeleton. Human Biol., *27*: 12-21.

RAO, C. R. 1952 Advanced Statistical Methods in Biometric Research. Wiley and Sons, New York.

SAUTER, M. R. AND F. PRIVAT 1955 Sur un nouveau procede metrique de

determination sexuelle du bassin osseux. Bull. Schweiz. Gesell. Anthrop. und Ethnol., 31 Jahrgang.

SCHULTZ, A. H. 1930 The skeleton of the trunk and limbs of higher primates. Human Biol., *2*: 303-438.

——— 1937 Proportions, variability, and asymmetries of the long bones of the limbs and the clavicles in man and apes. Human Biol., *9*: 281-328.

——— 1949 Sex differences in the pelves of primates. Am. J. Phys. Anthrop., n. s. 7: 401-423.

SIMPSON, G. G. 1941 Large Pleistocene Felines of North America. Amer. Museum Novitates. No. 1136.

STEWART, T. D. 1954 Sex determination of the skeleton by guess and by measurements. Am. J. Phys. Anthrop., n. s. *12*: 385-392.

THIEME, F. P. In press. Sex in Negro skeletons. J. Forensic Med.

TROTTER, MILDRED AND G. C. GLESER 1952 Estimation of stature from long bones of American Whites and Negroes. Am. J. Phys. Anthrop., n. s. *10*: 463-514.

WASHBURN, S. L. 1942 Skeletal proportions of adult langurs and macaques. Human Biol., *14*: 444-472.

——— 1948 Sex differences in the pubic bone. Am. J. Phys. Anthrop., n. s. *6*: 199-208.

——— 1949 Sex differences in the pubic bone of Bantu and Bushman. Am. J. Phys. Anthrop., n. s. 7: 425-432.

THE PHYSIOLOGICAL EFFECTS OF SLEEPING ON HARD OR SOFT BEDS[1]

BY E. E. SUCKLING, E. H. KOENIG, B. F. HOFFMAN AND CHANDLER McC. BROOKS

Department of Physiology, State University of New York Downstate Medical Center, Brooklyn, N. Y.

and

The United States Testing Co., Hoboken, N. J.

INTRODUCTION

CIVILIZED people employ a range of sleeping surfaces which includes those which are extremely soft and those which are very hard, often being but a wooden floor covered by a thin mat. It thus appears that man is able to sleep on any surface to which he is accustomed. The question of the advantage of one type of supporting surface over another appears, however, to be unsettled. With this problem in mind the present study has been conducted to determine whether there is any appreciable variation in certain physiological reactions recorded during sleep which can be related to the properties of the surface on which the sleeping subject is supported. Since it is not feasible to attempt to evaluate the quality of sleep under different conditions, sleep has been measured quantitatively in terms of its depth and duration. Quantification has been made on the basis of EEG recordings confirmed by stimulus-response studies. This objective evaluation has been compared to the nightly subjective estimate of the sleepers. In addition several other physiological variables (body movement, heart rate and skin temperature) have been recorded to supply still more information concerning the course of sleep during the night.

[1] This work was conducted in the Sleep Laboratory of the U. S. Testing Company. Support was received from the Simmons Company.

METHODS

The results discussed here were obtained from two adult male subjects who slept under observation for 5 nights a week during a period of 6 weeks and two others who slept under observation 5 nights a week for three weeks. This comprised a total of 90 sleep nights. The sleepers rotated through three different, constant temperature, sound-conditioned sleep rooms which differed only in the nature of the sleeping surface provided. One room contained a standard modern innerspring mattress, one a feather bed, and one a sheet of 1½ cm plywood covered by a thin carpet. The hardness of the various supporting surfaces was evaluated

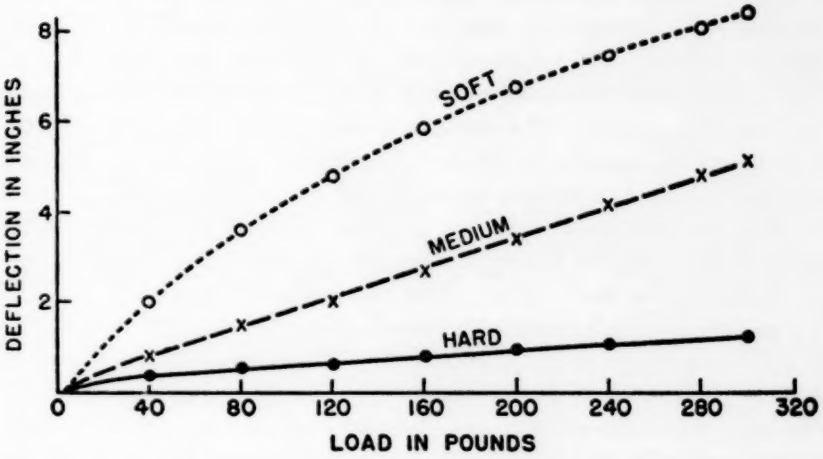

FIG. 1. LOAD-DEFLECTION CURVES FOR THE THREE SUPPORTING SURFACES.
○-----○ SOFT; ×---× MEDIUM; ●——● HARD.

by means of a deflection test. Known weights ranging up to 150 kg were placed on a board measuring 45 by 90 cm and located in the approximate center of the supporting surface. The depression of the board due to the compliance of the supporting surface was then registered over the range of weights used. Figure 1 shows that a marked difference in hardness was provided by the three supporting surfaces employed. In subsequent sections the board is referred to as the "hard" surface, the standard mattress as "medium," and the feather mattress as "soft."

The EEG was recorded from frontal, temporal and occipital leads with small silver electrodes sealed to the scalp by several layers of

collodion. A pre-sleep record was obtained and then immediately following lights-out a continuous record was taken for 20 minutes. In most instances this interval was sufficient to cover the initial going-to-sleep period of each subject. Subsequently the EEG was recorded for two minutes out of each 5 during the remainder of the 7-hour sleep period. The EEG records were evaluated by means of the usual criteria (Blake *et al.*, '39; Davis *et al.*, '39 and Gibbs and Gibbs, '50) and estimated depth of sleep was graded from 0 = awake to 4 = deepest sleep. (Hoffman *et al.*, '56 and Brooks *et al.*, '56).

The accuracy of the estimation of sleep-depth for each subject was checked by means of stimulus-response studies employing sound stimuli of known relative intensities. Tone pulses at 1500 cps and of 500 msec duration were delivered to each subject at widely spaced intervals on different nights. During these tests the EEG was monitored continuously and at each estimated depth of sleep, sound stimuli of increasing intensity were applied. The response to those stimuli was evaluated both in terms of the appearance of the "K complex" and also the stimulus intensity required to produce frank arousal of the subject.

Body movement was recorded by means of tambours attached from below to the center of the supporting surface. An attempt to equalize the recorders was made by adjusting the recorders' sensitivity for each bed so that standard movements gave deflections of similar magnitude on all three surfaces. The heart rate was determined from EEG records obtained simultaneously with the EEG. The electrocardiogram leads were attached to the anterior and posterior surfaces of the chest to avoid interference with limb movement. Skin temperature was recorded by thermocouples sealed to the plantar surface of the great toe and to the inner surface of the axilla. Each thermocouple was covered by cotton, silver foil, and several layers of tape. All electrode leads consisted of thin, flexible wires plaited into a soft cable which emerged from the pajama collar. This cable did not interfere with ordinary body movements during the night.

The subjective estimate of sleep was recorded in the following manner. Each subject was required to evaluate the previous night's sleep after wakening, by marking an appropriate point on a vertical scale ranging from 0 to 100. The range of variation in score for the 4 subjects differed widely.

All sleepers were engaged in their usual occupations during each day of the study. They reported to the sleeping rooms at 9 P. M. each night

and filled in questionnaires on the activities of the day, food intake, etc. The first three nights for each subject were considered as a breaking-in period and results from these nights were not used in subsequent evaluations.

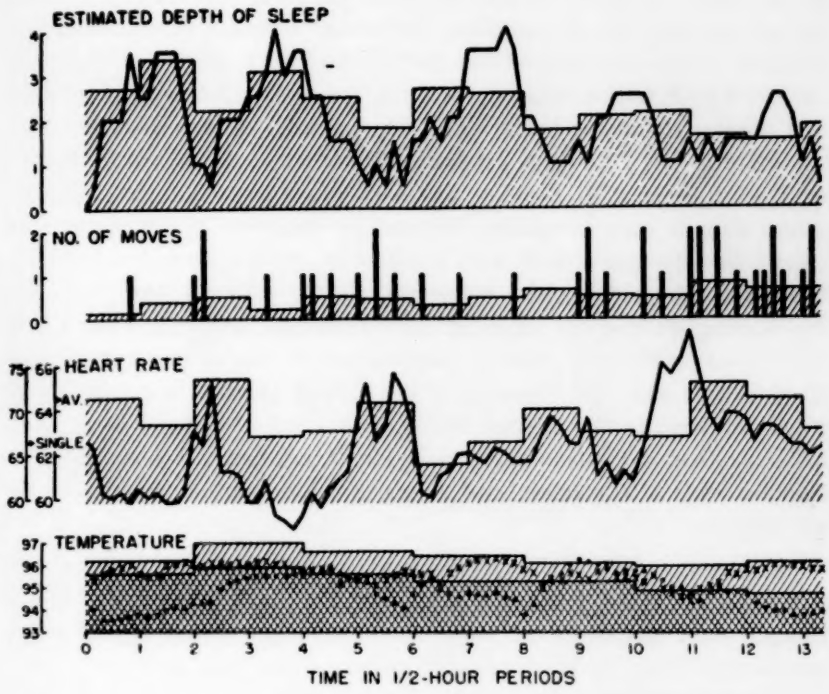

FIG. 2. COMPARISON OF SINGLE NIGHT AND AVERAGE VALUES FOR ONE SUBJECT FOR ALL SURFACES. SINGLE NIGHT PLOTTED AS TWO-MINUTE AVERAGES WITH HEAVY LINE, AVERAGE VALUES PLOTTED AS HALF-HOUR AVERAGES BY DIAGONALLY LINED AREA. TOE TEMPERATURE SHOWN WITH (. . .) HALF-HOUR AVERAGES (DOUBLE DIAGONALS), AXILLARY TEMPERATURE WITH (×××), HALF-HOUR AVERAGES (DIAGONAL LINES).

RESULTS

Figure 2 shows a sample single night record for one subject of the estimated depth of sleep, heart rate, frequency and duration of movement, and the toe and axillary temperature. Certain of these variables warrant some mention prior to evaluation of the effects upon them of the supporting surfaces. Most important of these is the change in EEG

pattern recorded at different periods of the night and its relation to the depth of unconsciousness.

Stimulus Testing of Depth of Sleep: When the EEG tracing showed predominant alpha rhythm (stage 0) or a low voltage pattern (null phase, stage 1) sound stimuli of moderate intensity did not elicit a typical and complete K complex. Somewhat stronger stimuli, however, gave rise to a wakening reaction characterized by a change in the EEG, an increase in muscle action potentials, frequent overt body movement, and often a change in heart rate. At somewhat deeper levels of sleep, identified in the EEG by a predominance of waves at 4-7 cps (stage 2) or the appearance of spindles at 14-16 cps (stage 3), low intensity sound stimuli were invariably followed by complete K complexes. At these levels stronger stimuli were required to produce arousal. In deepest sleep, characterized by delta waves and slow spindle activity (stage 4), the K complex was not recorded after moderate to strong stimuli and very intense stimulation was necessary to produce an awakening reaction. At this time also, the duration of the arousal reaction was often brief when compared to that elicited during lighter stages of sleep. In general, although there was some slight individual variation the same relationship between EEG pattern and response to stimulation was found for all subjects.

Electroencephalographic Evaluation: Direct observation of the subject and the effects of sound stimuli were both relied upon until it was thought that the depth of sleep for any subject could be estimated satisfactory from his EEG tracings alone. After this, each two-minute sample of the EEG was evaluated and assigned the appropriate numerical value (0-4). From these values a curve was constructed which represented the change in depth of sleep throughout the night (see figure 2). As has been reported previously (Hoffman *et al.*, '56; Brooks *et al.*, '56 and Kleitman, '39), there are a number of cyclic fluctuations in such a curve. The usual pattern shows a fairly rapid transition to deepest sleep (stage 4) within the first 60-80 minutes, and then a partial arousal. Subsequently, sleep again deepens but rarely reaches the depth attained initially. During the night the average levels and also the peaks of the curves tend to move from deeper to lighter sleep.

It was of some interest to note that for each subject, although there was considerable variation in the configuration of the curve on different nights, the cyclic nature of sleep depth was sufficiently consistent to

appear quite prominent in the average curve for any one sleep surface (figure 3) and even for all nights combined (figure 2). The consistency of pattern was most prominent during the initial period of going to sleep; this can be seen in figure 4A, which shows the frequency with which each depth of sleep occurred in each two-minute sample during

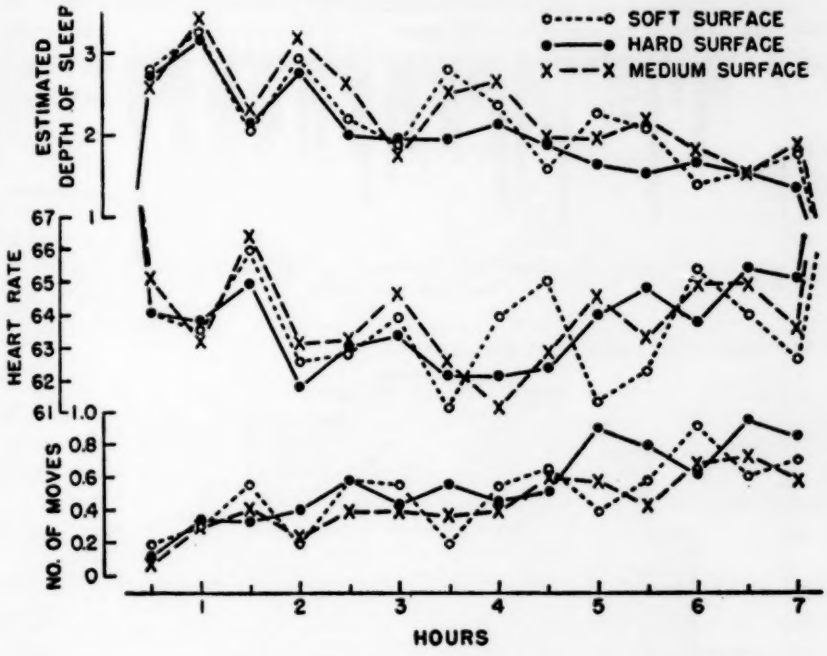

FIG. 3. AVERAGE VALUES PER HALF-HOUR FOR ESTIMATED DEPTH OF SLEEP, FREQUENCY OF MOVES, AND HEART RATE FOR SUBJECT "W" ON THE SOFT (○), MEDIUM (×), AND HARD (●) SURFACES. EACH CURVE REPRESENTS 10 NIGHTS AND EACH POINT THE AVERAGE OF 60 5-MINUTE PERIOD VALUES. HEART RATE EXPRESSED AS BEATS PER MINUTE, MOVES AS NUMBER PER 5-MINUTE PERIOD.

the first 80 minutes of the night. Similarly, when the average value for each half hour of the night is plotted in the same manner (figure 4B) the relative regularity of the fluctuations in sleep depth again becomes evident.

Heart Rate: The relationship between variations in heart rate and fluctuations in the depth of sleep are shown in figure 2. The reciprocal nature of this relationship has been demonstrated previously (Brooks *et al.*, '56) and it has also been shown that in spite of the good moment-

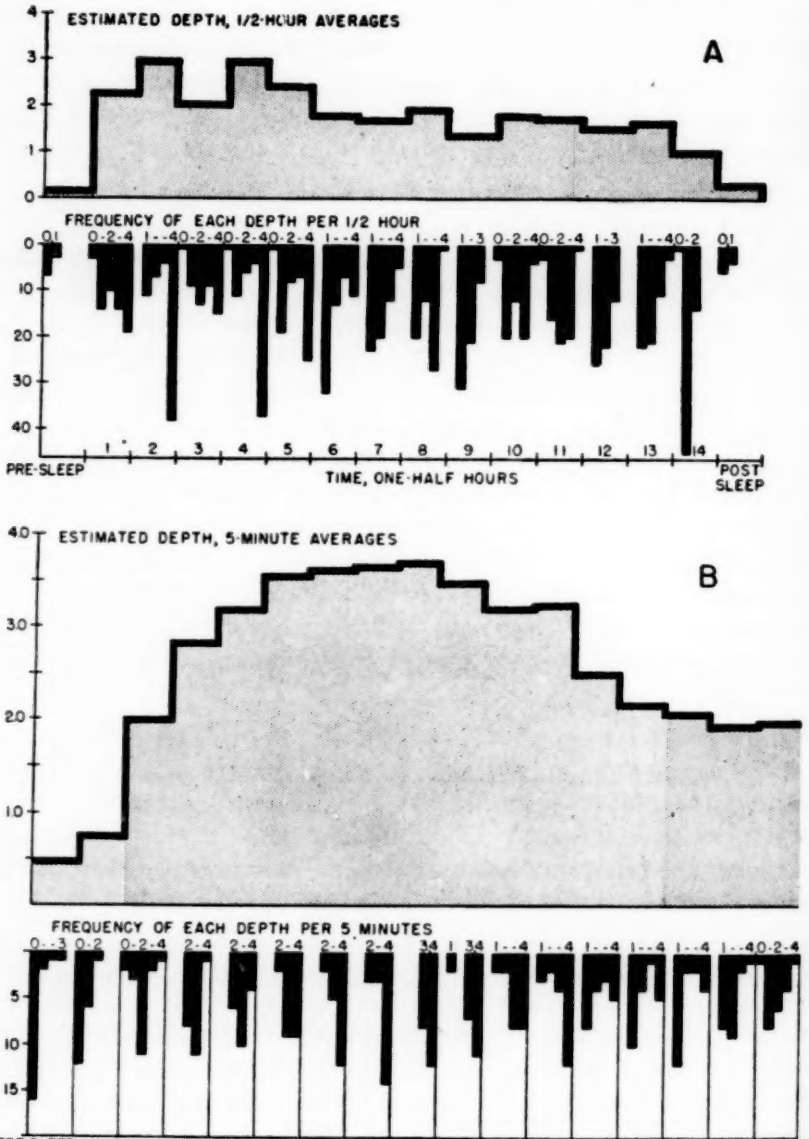

FIG. 4. A. AVERAGE ESTIMATED DEPTH OF SLEEP PER HALF-HOUR FOR ONE SUBJECT OVER 20 NIGHTS (TOP) AND FREQUENCY WITH WHICH EACH DEPTH OCCURRED DURING EACH HALF-HOUR. (BOTTOM)

B. AVERAGE ESTIMATED DEPTH OF SLEEP PER 5-MINUTE PERIOD DURING FIRST 80 MINUTES OF GRAPH IN A., AND FREQUENCY WITH WHICH EACH DEPTH OCCURRED DURING THE 5-MINUTE PERIOD.

to-moment correlation between estimated depth of sleep and heart rate, the *average* nightly heart rate is only poorly correlated with the average nightly depth of sleep (Brooks *et al.*, '56). Similar results have been obtained in the present study. Of some interest, however, is the relative magnitude of the change in heart rate for a certain change in the depth of sleep (figure 5A). Based on the average values for all subjects when the depth of sleep increased or decreased by two or three levels the heart rate slowed or accelerated by only 4 to 6 beats per minute. Since the average heart rate for all subjects was 62.5, this change amounts to

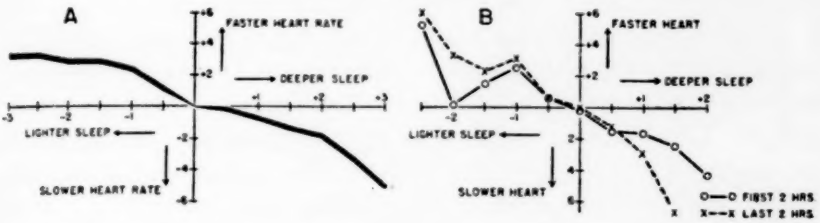

FIG. 5. A. AVERAGE CHANGE IN HEART RATE FOR A CHANGE OF GIVEN MAGNITUDE IN ESTIMATED DEPTH OF SLEEP BASED ON ALL VALUES FOR THREE SUBJECTS. ORDINATES—HEART RATE, BEATS/MINUTE; ABSCISSAE, CHANGE IN DEPTH OF SLEEP, STAGES 0-4.

B. AVERAGE VALUES FOR ONE SUBJECT, SHOWING COMPARISON BETWEEN FIRST TWO (O–O) AND LAST TWO (×–×) HOURS OF THE NIGHT.

less than 10%. Furthermore, the shape of the curves in figure 5B shows that the heart rate would be a less sensitive indicator of large alterations in sleep depth than of smaller changes in depth. Finally, the magnitude of the alteration in heart rate is not constant throughout the night; as seen in figure 5B, the slope of the line for the first two hours of sleep is significantly less ($p < .01$) than for the last two hours of the night.

Body Movement and Skin Temperature: Body movement appears to be associated with a decrease in the depth of sleep (figure 2). However, as has been shown in previous studies (Hoffman *et al.*, '56; Brooks *et al.*, '56 and Kleitman, '39), there is no consistent significant correlation between the number of recorded body movements on a given night and the depth of sleep on that night. The only other physiological variable consistently related in a significant fashion to frequency of body movements is the duration of each movement. Two other points concerning body movement became apparent in this study; during the first hour

of sleep, body movement is less likely to be associated with an appreciable decrease in depth of sleep than later during the night; second, there is a consistent increase in the frequency of movement of all subjects as the night progresses.

The results obtained from records of toe and axillary temperatures were not rewarding. Under the conditions imposed, there was no consistent progressive change in either temperature during the course of the night (see figure 2), nor was there any significant nightly average relationship between toe and axillary temperatures or between either temperature and any of the other variables studied.

Comparison of the Supporting Surfaces: Some impression concerning the effect of the different supporting surfaces on depth of sleep estimated from the EEG may be obtained from inspection of the average curves in figure 3. In order to obtain a basis for quantitative comparison the average depth of sleep per test period for each surface was determined. These values for each of the 4 subjects are shown in table 1. It is apparent that the magnitude of the difference in average depth of sleep on each of the three test surfaces differs quite markedly between the 4 subjects: for two of them (W and He), in terms of the area under the depth of sleep curve for the entire night, this difference is appreciable; for one (P), however, the values are practically identical. These differences for each subject on each surface are summarized in table 1. It can be seen that there is a significant difference for subjects W and He between depth of sleep on the medium and hard surfaces, and only for subject W in the case of the other two comparisons.

When the results obtained for all 4 subjects are averaged (table 1) it appears that there is a group preference for both the medium and soft over the hard surface. Similar results are obtained when the depth of sleep during either the first or second half of the night is employed in place of the average for the entire night. On the other hand, if the comparison is restricted to the first hour of sleep there is no appreciable difference between any of the test surfaces with respect to either the results for each individual subject or the average results of the entire group (figure 6).

It is interesting to compare the results obtained from estimation of the depth of sleep from the EEG with the evaluation of the goodness of their night's sleep on the different surfaces by each subject (table 2). It should be noted in observing the results that the 4 subjects rated

TABLE 1

Estimated depth of sleep

SUBJECT	AVERAGE VALUE PER SURFACE*		DIFFERENCE BETWEEN SURFACES Difference*	t	$p <$	Significant	Preference
	S	1.94	S–H = .05	.519	.7	No	—
Ho	H	1.89	H–M = —.11	1.170	.3	No	—
	M	2.00	S–M = —.06	.617	.6	No	—
	S	2.20	S–H = .18	2.67	.02	Yes	S
W	H	2.02	H–M = —.33	5.00	.001	Yes	M
	M	2.35	S–M = —.15	2.57	.02	Yes	M
	S	1.85	S–H = .22	1.57	.2	No	—
P	H	1.84	H–M = —.25	3.55	.01	Yes	M
	M	1.84	S–M = —.03	.30	.8	No	—
	S	2.26	S–H = .01	.096	1.0	No	—
He	H	2.04	H–M = 0.	0.	1.0	No	—
	M	2.29	S–M = .01	.124	1.0	No	—
	S	2.07	S–H = .117			Yes	S
Average of	H	1.95	H–M = .192			Yes	M
all subjects	M	2.14	S–M = —.075			No	—

* S = soft, H = hard, M = medium surfaces

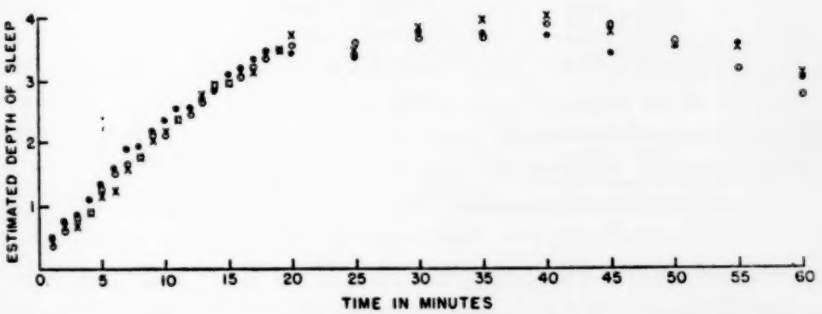

FIG. 6. AVERAGE ESTIMATED DEPTH OF SLEEP OF ALL 4 SUBJECTS ON EACH OF THE SUPPORTING SURFACES DURING THE FIRST HOUR OF THE SLEEP PERIOD. ○○○ = *soft*; ××× = *medium*; ●●● = *hard surface.*

their sleep on the three surfaces in widely different ways; thus the range of scores for one subject rarely varied by more than 5 units, while for the others the scores on different nights differed quite markedly. However, the data obtained do show that the subjects were aware of differences in the supporting surfaces and expressed quite definite opinions concerning the effect of the surfaces on their sleep. Thus all of the subjects rated the hard surface lowest of the three possible choices (table 2), and the 4 subjects treated as a group, showed a significant preference for either the soft or medium surface over the hard (table 2). On the

TABLE 2

Subjective estimate of sleep

SUBJECT	AVERAGE VALUE PER SURFACE*		DIFFERENCE BETWEEN SURFACES Difference*		t	$p <$	Significant	Preference
	S	85.6	S–H =	30.5	3.66	.01	Yes	S
Ho	H	55.1	H–M =	—29.1	3.79	.01	Yes	M
	M	84.2	S–M =	1.4	.348	.8	No	—
	S	62.9	S–H =	1.0	.297	.8	No	—
W	H	61.9	H–M =	—3.1	.881	.4	No	—
	M	65.0	S–M =	—2.1	.612	.6	No	—
	S	97.4	S–H =	27.2	8.69	.001	Yes	S
P	H	70.2	H–M =	—27.6	8.88	.001	Yes	M
	M	97.8	S–M =	—.4	.89	.4	No	—
	S	63.6	S–H =	4.2	.383	.8	No	—
He	H	59.4	H–M =	—28.6	3.06	.02	Yes	M
	M	88.0	S–M =	—24.4	3.75	.01	Yes	M
	S	76.2	S–H =	15.9			Yes	S
Average of	H	60.3	H–M =	19.0			Yes	M
all subjects	M	79.3	S–M =	3.1			No	—

* S = soft, H = hard, M = medium surfaces

other hand, when the subjective estimates of the individual sleepers are analyzed (table 2) it can be seen that there is no consistent relationship between the statistically significant preferences based on the estimated depth of sleep and those based on the subjective estimate of the sleeper. This failure of correlation is probably due to differences in basis of evaluation and should not bring into question the basic claims of objective measurement.

A somewhat similar picture is presented by analysis of the movement of the various sleepers during the night. In all cases the average number of moves was greatest on the hard surface (table 3) and results

TABLE 3

Number of moves/5 minute period

SUBJECT	AVERAGE VALUE PER SURFACE*		DIFFERENCE BETWEEN SURFACES Difference*		t	$p <$	Significant	Preference
	S	.602	S–H = —.217		1.75	.1	No	—
Ho	H	.819	H–M =	.263	2.22	.05	Yes	M
	M	.556	S–M =	.046	.417	.7	No	—
	S	.498	S–H = —.061		1.17	.3	No	—
W	H	.559	H–M =	.117	2.16	.05	Yes	M
	M	.442	S–M =	.056	.919	.4	No	—
	S	.258	S–H = —.283		6.79	.001	Yes	S
P	H	.541	H–M =	.208	3.76	.01	Yes	M
	M	.333	S–M = —.075		1.70	.2	No	—
	S	.536	S–H = —.006		.100	1.0	No	—
He	H	.542	H–M =	.174	3.33	.02	Yes	M
	M	.438	S–M =	.168	3.54	.01	Yes	M
	S	.501	S–H =	.147			Yes	S
Average of	H	.648	H–M =	.208			Yes	M
all subjects	M	.444	S–M =	.057			No	—

* S = soft, H = hard, M = medium surfaces

for the group of subjects revealed a real preference for either the soft or medium over the hard surfaces (table 3). However, as in the case of the subjective estimate, there was no consistent correlation between the significant results for individual sleepers based on movement and those based on estimated depth of sleep or subjective judgement.

The results of studies of average heart rate are shown in table 4. As might be expected from the results described above, the heart rate was practically identical on all surfaces for each of the 4 subjects. Temperature studies were similarly unrevealing, the only positive result being a suggestion that this value was higher when the sleeper was on the soft surface. If the soft surface tended to envelop the extremities more than the other two surfaces, such a result would be expected. Only one other point should be mentioned. In spite of the apparently

completely unyielding nature of the hard surfaces, none of the subjects made any strong complaints about spending one night out of three on a plywood board and none suggested that such a sleeping surface had an adverse effect on ability to perform usual daily tasks. Whether or not different results would have been obtained if the subject had been required to spend a large number of consecutive nights on this surface cannot be decided on the basis of these studies.

TABLE 4

Average heart rate, beats/minute

	SUPPORTING SURFACE		
SUBJECT	Soft	Hard	Medium
Ho	65.8	66.5	66.1
W	63.3	63.5	63.6
He	57.6	56.9	56.9
P	59.4	57.8	57.6
Average of all subjects	62.6	62.4	62.4

DISCUSSION

The results obtained in this study of the effect of three different supporting surfaces on several aspects of sleep bring up a number of questions which require additional elucidation. In the first place, while it is evident from the results of estimated depth of sleep, subjective estimate and body movement that the various surfaces did influence the character of the sleep obtained, it was not possible to evaluate the importance of the apparently small differences noted. In time small differences might create more obvious advantages. The change in average depth of sleep, if multiplied by the duration of the sleep period, indicates that at least for some individuals there was a substantial difference in the total quantity (depth times duration) of sleep on certain surfaces. It is not known whether this change in quantity of sleep has any significance in terms of the overall well-being of the individual. The poor correlation between subjective estimate and estimated depth of sleep might imply that it has none. Similarly, although there was a reasonable difference in the number of moves on the different surfaces, it appeared in this study, as in previous work (Hoffman *et al.*, '56; Brooks *et al.*, '56 and Kleitman, '39), that there was little relationship

between average depth of sleep and number of moves during the night. Although there was a somewhat more consistent relationship between movements and subjective estimate than between movements and sleep depth, even in the former case it would appear that the sleepers' estimate is influenced strongly by factors other than body movement during sleep. The consistency of the heart rate under all three test conditions is not surprising in view of the fact that it is only an insensitive index of changes in depth of sleep. Finally, it should be emphasized that there is no information available concerning the actual advantages to the sleeper of a higher or lower value for any of the variables studied. In this sense, then, no conclusions concerning the effect of the supporting surfaces on the *quality* of sleep can be attempted. However, there are two positive results of this study. In the first place, it has been demonstrated that individuals differ considerably in their sensitivity to supporting surfaces. Secondly, quantitative differences were seen and though no conclusions concerning quality and beneficial effect can be substantiated at this time, it can be said that some of the subjects definitely slept less well on the very hard surface and one slept less well on both the hard and soft, than on a medium supporting surface. In certain circumstances these small advantages and disadvantages might be magnified.

SUMMARY

The EEG, heart rate, body movement, and skin temperature of 4 subjects sleeping on three different (hard, medium, and soft) supporting surfaces have been studied and the subjective estimates of the sleepers recorded. The use of sound stimuli has been found of considerable value as an aid to the identification of the depth of sleep from the EEG records. As in previous studies, a good inverse relationship has been found between moment to moment changes in heart rate and sleep depth. However, this relationship is not quantitatively the same throughout the night, and the average heart rate is not related to the average depth of sleep. The relationship between movement and depth of sleep is somewhat similar to that for heart rate. Factors other than depth of sleep are of importance to the control of these variables. On the basis of the studies performed, it can be said that the hard surface did have an effect on depth of sleep, movements, and the subjective estimate of sleep. However, the differences in these variables were not large and were not always statistically significant. Subjects varied in their sensitivity to changing firmness of supporting surfaces.

LITERATURE CITED

BLAKE, H., R. W. GERARD AND N. KLEITMAN 1939 Factors influencing brain potentials during sleep. J. Neurophysiol., *2*: 48-60.

DAVIS, H., P. A. DAVIS, A. LOOMIS, E. N. HARVEY AND G. HOBART 1939 Electrical reactions of the human brain to auditory stimulation during sleep. J. Neurophysiol., *2*: 500-514.

GIBBS, F. A., AND E. L. GIBBS 1950 Atlas of Electroencephalography. 2nd ed., 2 vols., Addison-Wesley Press, Inc., Cambridge, Massachusetts.

BROOKS, C. McC., B. F. HOFFMAN, E. E. SUCKLING AND F. KLEYNTJENS 1956 Sleep and variation in certain functional activities accompanying cyclic changes in deep sleep. J. App. Physiol., *9*: 97-104.

HOFFMAN, B. F., E. E. SUCKLING, C. McC. BROOKS, E. H. KOENIG, K. S. COLEMAN AND H. J. TREUMANN 1956 Quantitative evaluation of sleep. J. App. Physiol., *8*: 361-368.

KLEITMAN, N. 1939 Sleep and Wakefulness. University of Chicago Press, Chicago.

BOOK REVIEWS

Body-Build, Body-Function, and Personality. Edited by BENGT LINDEGÅRD. 132 pp. Lunds Universitets Årsskrift. N. F. Avd. 2. Bd 52. Nr 4. 1956. No price indicated.

The editor, who is also the author or co-author of all the 6 papers making up the volume, is in charge of the Laboratory of Differential Somatology, associated with the Department of Anatomy, University of Lund, Sweden. For some time he has been carrying on, industriously and with imagination, studies on individual differences. As the subtitle of the volume indicates, these are reports on different aspects of "Medical anthropological investigations on 320 Swedish 20-year-old healthy army men." This defines the sample. How about the content?

The system for a quantitative assessment of body-build, described in detail in an earlier monograph,* is briefly presented in the first paper ("Differential Morphology"). It operates with variables chosen as to reflect individual variations in (1) body length (length of radius and tibia), (2) skeletal sturdiness (condylar breadths—femoral, tibial, bimalleolar), (3) muscular development (from dynamometric data on grip strength, shoulder thrust, and shoulder pull), and (4) adiposity (evaluated as a "residual," on the basis of body weight or extremital girths, keeping other relevant body-build "factors" constant, or from skinfolds).

A study, concerned with fat distribution and utilizing skinfolds measured at 10 sites, suggests the presence of a continuous interindividual variation as regards the proportion between truncal and extremital deposits of subcutaneous fat (B. Škerlj's Truncic-Extremital Vector; Am. J. Phys. Anthropol. (n.s.), 11, p. 591, 1953). Muscularity, assessed by dynamometric recordings, was examined in reference to physical activity and selected male sex characters, such as genital size and hairiness of the trunk, were examined—with some unexpected results, such as hairiness being positively related to fatness—in reference to general body-build, endocrine activity, and personality. The relationship between some steroid substances excreted in the urine of adult

* Variations in Human Body-Build. By Bengt Lindegård. Acta Psychiatr. Neurol., Suppl. 96. E. Munksgaard, Copenhagen, 1953.

males were subjected to biometric analysis and the interrelations between somatologic, endocrine and psychological variables were studied in detail.

Space does not permit a detailed critical appraisal of the author's system of *differential somatology* and his exploration of the co-variation with other somatologic, biochemical, and behavioral characteristics. The interdisciplinary approach, involving the collaboration of endocrinologists (S. Genell, C. C. Jensen) and students of personality (G. E. Nyman, a psychiatrist) is in the best tradition of human biology. The system of presenting the basic variables ("factors") in the form of a "body-build diagram," with group means as reference points and standard deviations as units of measurement, while firmly established in psychology, in the form of ability or personality "profiles," is new in physical anthropology (*cf.* S. M. Garn, Ann. N. Y. Acad. Sci., *63*: 537-552, 1955). Lindegård's search for a more meaningful expression of somatic variables and his use of "residuals," *i. e.*, differences between the actual value of a measurement, such as body weight, and the value predicted from the correlated dimensions, such as those indicative of body length, deserves notice.

This does not mean, however, that the proposed system should be swallowed in its totality—hook, line, and sinker—by anthropometrists still hungry for a workable system of quantitative description of human physique or that this reviewer is celestially happy with all other aspects of the work by Lindegård and his co-workers, including the assessment of personality in terms of four variables, labelled "capacity," "validity," "stability," and "solidity," and rated on a 7-point scale on the basis of short, single interviews.

It is unfortunate that Lindegård was not familiar with the far-reaching ideas of J. Matiegka (Am. J. Phys. Anthrop., *4*: 223-230, 1921) who was motivated by that same healthy discontent with the existing systems for the characterization of human physique and who presented a scheme for describing body build in terms of relative amounts of muscle and adipose tissue, with skeletal weight as the reference point. Quantitatively oriented anatomists owe Matiegka a debt, still largely unpaid (except for Mildred Trotter's work on the estimation of skeletal mass; Am. J. Phys. Anthrop., (n. s), *12*: 537-551, 1954), of replacing Matiegka's approximative constants by rigorously derived estimation equations, especially of the total amounts of muscle in the body.

The combination of performance criteria used in assessing "muscu-

larity" with morphological criteria, which are the basis for a quantitative appraisal of skeletal "length" and "sturdiness," is undesirable. It must be recalled that performance, including voluntary muscle contractions, is a "product" of capacity (a physiological concept, related to muscle size) and motivation. Only when the latter is "constant," preferably maximal, can actual performance serve as indicator of capacity and, in turn, of muscular development. This condition was probably fulfilled in Lindegård's series but can not be postulated as being present in all situations. Furthermore, such a mixture of approaches is not necessary. Measurements of the circumferences of the extremities together with skinfolds determined at the same level (with femoral condyles and a length factor added for good measure) provide an adequate indicator of muscular development, at least for the extremities.

The separate pagination of individual papers, and consequent absence of a table of contents or an index, does not make for efficient use of this valuable and stimulating volume.

JOSEF BROŽEK

University of Minnesota,
Minneapolis, Minn.

Radiographic Atlas of Skeletal Development of the Knee: a Standard of Reference. By S. IDELL PYLE and NORMAND L. HOERR. 82 pp. Charles C. Thomas, Springfield, Illinois, 1955. $4.25.

The heaviest single bone in the human body is the femur. Most of its elongation occurs at the distal end, at an epiphysial site which can readily be studied in children from radiographs of the knee. The atlas of Pyle and Hoerr now enables us to trace the skeletal development of this region from infancy to maturity.

This atlas is based on nearly 30 years of research on developmental radiographs of the skeleton which have been accumulating at Western Reserve University. These monumental studies were initiated by the late Professor T. Wingate Todd. It was he who directed the growth study of the Brush Foundation in Cleveland, Ohio through its first decade, and published his *Atlas of Skeletal Maturation* (*Hand*) just before his death in 1938. This was an interim atlas, based largely on cross-sectional surveys of Cleveland children from both poor and prosperous environments. Todd's successor, Dr. Walter W. Greulich, con-

tinued the direction of the Brush Foundation study. In 1950, he and Dr. S. I. Pyle produced a more definitive atlas of the hand. The Greulich-Pyle atlas is now used for developmental assessments in many parts of the world. It is based on selected Cleveland children, reared under economically favored conditions and in optimal health. These children were relatively free from the stresses of malnutrition, illness and climate which impede the growth of much of mankind, and the Greulich-Pyle standards therefore present an idealized picture of rapid and regular skeletal development.

Nearly all of the films in the Pyle-Hoerr atlas of the knee are also based on children from the Brush Foundation study, with optimal histories of health. The book contains 57 films of the knee, which were very carefully culled from an original series of 10,700. Some of the children were followed for as little as two years, and others for as long as 12.

The Greulich-Pyle atlas of the hand has separate sets of standards for boys and girls, but the Pyle-Hoerr atlas has a single set of films of the knee for both sexes. Girls undergo more rapid skeletal maturation than do boys however, so that after the first few postnatal weeks, each film has a younger developmental age equivalent for girls than for boys. Up to 6 years of age, and from 10 to 12 years, an average girl at a given stage of ossification of the knee is about 80% as old as a corresponding boy. From 7 to 9 years, she is at about 70% of the boy's age. Prior to adolescence, these percentages may be of use in estimating the age and sex of immature skeletal remains from bones and teeth, as outlined in 1955 by Hunt and Gleiser (Am. J. Phys. Anthrop., n.s. *13*:479-488).

In practice, the problem arises as to whether skeletal age should be assessed from both hand and knee in an individual child. For the hand, one film is enough for a single assessment, while for the knee at most ages it is desirable to have both a mediolateral and a postero-anterior film. In routine work, the knee usually fails to add much to what a careful study of the hand can tell us. But if the clinician or investigator has special reasons for studying the developmental normality of the lower limb, he should certainly use the Pyle-Hoerr atlas. It is a product of years of devoted labor, and is deserving of high praise.

EDWARD E. HUNT, JR.

Forsyth Dental Infirmary for Children,
Boston, Mass.

Psychology, Evolution and Sex. By CECIL P. MARTIN. xiii + 166 pp. Charles C. Thomas, Springfield, Illinois. 1956. $4.75.

Professor Cecil P. Martin, of McGill University, is an anatomist who has grave suspicions about conventional genetic theory. Once before he issued a minority report (Martin '53), and now he does it again, in a book that has the most eye-catching title in recent memory.

Geneticists, says Dr. Martin, worship the orthodoxy of the "mutation-selection" theory, explaining away all that does not square with it, yet rejecting evidence to the contrary, and ignoring theories not contained within the several ideas of particularate inheritance, isolation of the genic material from environmental influences, natural selection, and mutation. Not quite a Lamarkian, and too much a naturalist to offer an *élan vital*, Martin assembles citations to prove "lingering modifications" (*Dauermodifikation*) and offers his own theory of why sex is important. Throughout, the argument is directed against what he thinks geneticists think, and that is why the objects of the objections are often so elusive.

Now the study of evolution is an evolving science where the heterodoxies of one era become the orthodoxies of another. The early post-Darwinian age of acceptance gave rise to the era of scepticism, and this in turn yielded to the *Genetical Theory of Natural Selection*. New ideas are to be welcomed; we still have much to learn about evolution. But Dr. Martin's ideas are a curious blend of last-century objections and half-understood concepts, both better illustrated than described.

Mutations, he says, are pathological conditions. Why? Because known mutagens, such as nitrogen mustard and penetrating radiation are exceedingly injurious. What happens, he argues, is that the genes or "particles of the genes" are agitated, much as violent agitation can dislocate a man's thumb. And from this line of anthropomorphic reasoning, it is not difficult for him to decide that *all* mutations (being injuries) are harmful, whereas such characters as seem to be adaptive are simply examples of "lingering modifications." Ergo, as stated on page 81 ". . . I fail to see how the mutation-selection theory can be defended."

As promised, psychology is drawn in, that is, the psychology of what used to be called instinctive behavior. Martin correctly cites many instances where the feeding preferences of a form have been modified by rearing the young on a non-preferred plant, on an unaccustomed

diet. Such examples, he feels, show that preferences can be modified and the new preference then becomes hereditary. The rest of us would see here the effects of learning: increasingly much of what used to be called instinctive, now appears to be learned. Even such seemingly basic "instincts" as the nesting instinct in the pregnant rat, or the "grooming" instinct depend upon the previous experiences of the animal in question. So the whole argument about the modification—and—inheritance of behavior dissolves into a simple exercise in learning theory.

As to sex, Martin rejects the conventional explanations. The idea that new genotypes are thus produced is insufficient for him; "shuffling the genes destroys gene-complexes just as fast as it creates them." (This, of course, is not at all true.) But his explanation for the survival of sex involves "its usefulness as a restriction on the transmission of the prejudices, aversions and partialities of parents to their offspring." For an understanding of this explanation, the book must be read.

The real value of the present work lies in the large number of rarely-quoted minority reports he cites. They include the work of Kammerer, whose unfortunate history Martin does not recount, and other publications where modifications were claimed to persist for many generations. They include, also, accounts of the reversion of domesticated lines to "wild" types, and others where hereditary traits are shown to persist only on a particular diet. (They do not include references to DeLamarck, whose inferences that the peculiarities of domesticated animals and plants are due to nutrition, are at least partly correct.)

What *Psychology, Evolution and Sex* does show, is the increasing difficulty of tackling one field, with the conceptual tools and references of another. Despite Martin, the autopsy table is not a proving ground for genetical theory, nor does embryological similarity prove that no new evolutionary characters have entered into pre-natal development.

STANLEY M. GARN

Fels Research Institute,
Yellow Springs, Ohio

Anxiety and Stress: An Interdisciplinary Study of a Life Situation. By H. Basowitz, H. Persky, S. J. Korchin, and R. R. Grinker. xv + 320 pp. McGraw-Hill, New York, 1955. $8.00.

In 301 pages of text, figures, and tables, a longitudinal study of some biochemical and psychological responses of parachute trainees is presented in meticulous and overwhelming detail. The subjects are healthy, highly selected individuals, and are keenly motivated to acquire the status of a parachutist. These facts alone almost completely destroy the effectiveness of the research. The trainees simply are not made very anxious by the training program, although tower-jumping and airplane-jumping may evoke increments of anxiety. "Free anxiety," however, is not "released" in large amounts, and the low degree of anxiety evoked is related more to fear of failure than to fear of bodily harm. With an integrity and determination that leaves this reviewer awed, the Michael Reese group nevertheless persevered through the execution of a series of small-sample studies of "anxiety and stress." To their credit, they do wring out of a mass of data some suggestive hypotheses, and a few clear findings.

The psychological measures (anxiety questionnaire, elaborate ratings of interview materials, tachistoscopically administered Bender-Gestalt test, serial subtraction, perceptual closure of incompleted circles presented tachistoscopically, memory for digits, and scores on a story-telling test using presumably anxiety-provoking pictures relevant to parachute training) were selected a priori and empirically. The lack of clear-cut conceptualization and theoretical focus is unfortunate.

The biochemical aspects of the study stem from a previous finding that hippuric acid excretion is high in highly anxious patients. This variable is the focus of the biochemical research. The other variables measured (plasma glycine level, plasma amino acid level, blood reduced and oxidized glutathione levels, eosinophile counts, and hematocrit) were introduced either in an attempt to explore the precursors of hippuric acid excretion, or to add information about adrenocortical response.

Decisively positive findings are few. Change in glutathione level appears to be an important biochemical stress response, intermediate between the overly sensitive eosinophile response and the insensitive hippuric acid response. Of the psychological variables, perceptual adequacy showed promise, and sometimes effectiveness, as a correlate of rated anxiety and of the stressfulness of activity. The pioneering

study of intra-individual variability included in this report should stimulate research in this sadly neglected area.

This monograph, to the reviewer's mind, is a report for specialists only, who are interested in technical details, and who are willing to tolerate the lengthy presentation of many negative, or only suggestive and provocative, results. The non-specialist is likely to end the book with the correct conviction that biochemistry, psychology, and psychiatry, either independently, or in partnership, have not forged either tools or concepts able to cope with the complexities of the interrelated responses of the "total organism" to a "life situation." The book ends with these words: "The gain from the work has been less in increasing our knowledge of the details of psychosomatic problems and more in learning the complexities of psychosomatic organization and integration. The problem of anxiety is still *non liquet*, but perhaps we know a little more about why this is so."

JOHN I. LACEY

Fels Research Institute,
Yellow Springs, Ohio

Population Theory and Policy: Selected Readings. Edited by JOSEPH J. SPENGLER AND OTIS DUDLEY DUNCAN. x + 522 pp. The Free Press, Glencoe, Illinois, 1956. $7.50.

One of the problems of a field like demography, overlapping so many areas of the social and biological sciences, is the inevitable scattering of its literature through an assortment of unrelated journals. This makes it difficult for both the classroom student and the general reader to obtain convenient access to what is published. The volume under review is one of two that should help to repair the situation, though its contents will for the most part be more appropriate for advanced than for beginning readers in the subject.

The 38 articles, with 78 pages of footnotes, are organized into 9 chapters, each of which begins with a brief introduction by the editors. The chapters are: Development of population theory; The role of theory in population studies; Theory of population growth; Population, resources, technology, and levels of living; Population and level of economic activity; Population growth and economic development; Population growth and international relations; Socio-cultural context of

population dynamics; and Population policy. With a very few exceptions, the material has been chosen from the writing of American authors, about two-thirds of it from publications of the last 10 years. The volume closes with a 9-page bibliography of recent works in English. The companion volume, *Demographic Anaylsis: Selected Readings*, reviewed elsewhere in this journal, presents some of the specific findings of demographic research, but neither of the two books consider in detail the analytical techniques of the field.

The general reader whose major interest lies in biology will probably not wish to read the whole book, but he will find it interesting to look through the volume to gain an idea of some of the lines of thought now being explored in demography. He may be struck in particular by the weight given to economic aspects of the problems considered. This is a reflection of the fact that although demography is often thought of in this country as more nearly a branch of sociology than of any other discipline, a great deal of the theorizing in the field has been done by economists. Direct attention to biological topics is quite rare in this volume and consists in part of a history of dubious analogies. Demography has still some progress to make in the effective use of biological information.

A few of the articles may be mentioned as likely to be of especial interest to readers of this journal. The three articles in the first chapter on the development of population theory contain a number of references to ideas of biology. Material from a variety of specialized areas is imaginatively drawn together in Boulding's "Toward a general theory of growth." Lotka's "Some recent results in population analysis" illustrates a major branch of demographic activity, while Davis' "Population and power in the free world" has broad general interest.

The chapter on population theory leaves something to be desired. A general coverage of important points of view in this area should include, for example, an exposition of the Marxian position which must be taken into account (even when it is not accepted). A more comprehensive description and assessment of policy measures affecting population in various countries would also add to the completeness of the presentation.

GEORGE F. MAIR

Smith College,
Northampton, Massachusetts

Demographic Analysis: Selected Readings. Edited by JOSEPH J. SPENGLER AND OTIS DUDLEY DUNCAN. xiii + 819 pp. The Free Press, Glencoe, Illinois, 1956. $9.50.

This lengthy companion volume to *Population Theory and Policy* contains a total of no less than 63 articles selected from 25 different journals and other more inaccessible sources. Although there are a number of selections of particular relevance for the student of biology, the orientation of the collection is revealed by the predominance of sociologists and economists among the contributors.

The volume is admittedly not, as its title might suggest, designed to be a "do-it-yourself" guide to the methodology and analytic techniques of demography. The basic criterion for the inclusion of an article was presumably that it illustrate both the demographer's approach to population data and the types of research problems that are posed. The book is divided into 7 sections, each introduced by brief organizing editorial remarks. The broad range of subjects covered includes the conventional topics of past and future population trends, mortality, fertility, international and internal distribution of population and migration, and population composition. The volume concludes with a series of papers describing the recent demographic histories of over a dozen countries and areas throughout the world. The section of most direct interest to the biologist is Chapter VI (population composition), which includes 5 selections on population quality. The primary concern of these articles is with the presumed genetic consequences of population change, particularly differential fertility. In this context the subject of the future intelligence levels of the population is reviewed. These selections strikingly reveal the primitive level of measurement criteria in eugenics although they indicate no dearth of opinions.

An evaluation of a volume such as this is not simple. Certainly the value of each contribution cannot be appraised individually. The conventional standards that remain, therefore, are the selections included, the editorial skills evidenced in the introductory sections, and the potential usefulness of such a book. The strongest feature of the book is the high quality of the choices made by the editors. On the whole, the articles reprinted are excellent both in the sense of illustrating the methodological approaches of demography and of assembling in one source some of the most important findings of demographic research. Occasionally the reader will feel uncomfortable in the presence of

assumptions made about the inevitable stabilization or decline of the United States population contained in articles written over ten years ago when such assumptions were fashionable. However, it provides one with an interesting picture of the profession's self-consciousness as revealed in Harold Dorn's excellent critical review of population forecasting.

The editors' introductory materials are not sufficiently long to provide the kinds of coherence and organization that might be desirable. The value of the book rests squarely on their wise choice of materials rather than the connective tissues they superimposed.

The usefulness of the volume will, I believe, be quite limited owing both to its price and the fact that the materials are more suitable for the graduate student and professional than for the undergraduate. Although its usefulness to the biologist is even more limited, the book can be regarded as an excellent source book of demographic research classics.

CHARLES F. WESTOFF

Office of Population Research,
Princeton University

Non-parametric Statistics for the Behavioral Sciences. By SIDNEY SIEGEL. 312 pp. McGraw-Hill, New York, 1956. $6.50.

This book forms a valuable addition to the McGraw-Hill Series in Psychology because over half of the text presents detailed information about procedures not described in most of the popular statistics texts. It does so in a very clear manner, illustrating the computations for each procedure by a worked out example and will be useful to workers in many areas of human biology.

Although the remainder of the material may be found in some other books, it is presented here in a novel frame of reference. In fact, perhaps the most useful feature of the text is the strong emphasis placed throughout on the fundamental distinctions between procedures applicable to nominal, ordinal and interval scales. A table, printed on the inside of the front and back covers of the book summarizes the techniques discussed and forms a fine pedagogic device continually reinforcing these distinctions.

The text will be of great help to anyone who wishes to enrich his kit of statistical tools, and should present no difficulties to persons not well acquainted with mathematic statistics. As a textbook it will supplement the various available texts concerned with the better known parametric statistics such as t, F and r. A glossary of symbols, a combined author and subject index, a list of references for further reading and a large number of tables not readily accessible elsewhere enhance the usefulness of the book.

STEVEN G. VANDENBERG

University of Michigan,
Ann Arbor, Mich.

BOOK NOTICES

GENERAL SYSTEMS. Edited by Ludwig von Bertalanffy and Anatol Rapoport. 162 pp. Society for the Advancement of General Systems Theory, 1956. (no address, no price) The first yearbook of the Society consists of 3 parts: (1) Introduction (3 papers), (2) Exploration of mathematical models, with special reference to life processes (evolution, growth, individual and group behavior), and (3) "in search of new foci of integration" in biology, psychology and sociology.

FATIGUE, STRESS, BODILY CHANGE AND BEHAVIOR: SELECTED BIBLIOGRAPHY. By William Bevan and Rollin M. Patton. 59 pp. (no price) Military Operations Research Engineering Division, Lockheed Aircraft Corporation, Marietta, Georgia, 1956. An alphabetical listing of some 900 titles, followed by a topical index. The bibliography, covering the years 1946-56, was compiled as part of preparations for experimental investigations on aircrew fatigue.

MOLECULAR STRUCTURE AND FUNCTIONAL ACTIVITY OF NERVE CELLS. Edited by R. G. Grennell and L. J. Mullins. ix + 169 pp. American Institute of Biological Sciences, Washington, D. C., 1956. $4.75. The proceedings of a symposium held in Washington, D. C., on June 3-4, 1955, under the sponsorship of the Office of Naval Research dealing with 3 groups of subjects; (1) receptors (chemoreception, vision), (2) aggregates of neurons, and (3) nerve cell membranes. The publication demonstrates one way in which the Institute can contribute significantly to advancement in biological sciences.

HYPNOSIS AND ITS THERAPEUTIC APPLICATIONS. Edited by Roy M. Dorcus. x + 327 pp. McGraw-Hill, New York, 1956. $7.50. Aims to incorporate in a single volume the theoretical, experimental, and therapeutic aspects of hypnosis. Includes a large chapter on the physiological effects of hypnotic stimulation.

DICTIONARY OF ANTHROPOLOGY. By Charles Winick. vii + 579 pp. Philosophical Library, New York, 1956. $10.00. "A field which includes archeology, cultural anthropology, linguistics, and physical anthropology has boundaries and a vocabulary that are constantly expanding." Cautious souls will regard with uneasiness one man's attempt to master such a wide area. Physical anthropologists will find that the expanding edge of their science is not adequately represented.

PROGRESS IN THE PHYSIOLOGY OF FARM ANIMALS. Edited by John Hammond. v + 330 pp. Butterworths, London, 1957. 50s. ($7.00). The final volume of a comprehensive survey, not lacking wider implications (see *Human Biology*, 28: 278-280, 1956), deals with reproduction (reproduction hormones, female fertility, pregnancy) and milk production.

MEDICAL EFFECTS OF THE ATOMIC BOMB IN JAPAN. Edited by Ashley W. Oughterson. xvi + 477 pp. National Nuclear Energy Services, Manhattan Project Technical Section, Division VIII, Vol. 8. McGraw-Hill, New York, 1956. $8.00. Based on a 6-volume report, this book presents a summary of ecological, clinical, hematological, and pathological studies made in Hiroshima and Nagaski following the atomic bombing of the two cities, resulting in mechanical, thermal, and radiation injuries. A "tract for our times," indeed.

CULTURE AND THE STRUCTURAL EVOLUTION OF THE NEURAL SYSTEM. By Fred A. Mettler. 57 pp. American Museum of Natural History, New York, 1956. In this James Arthur Lecture on the evolution of the Human Brain (1955), the author has drawn attention "to the fact that there is no good evidence to support the assumption that man's neural system has undergone any progressive alteration in the direction of greater size or complexity since the middle of the Pleistocene. It was at this period that cultural development became manifest, and the hypothesis is advanced that, with the advent of culture, man achieved a degree of independence from his environment and was therefore no longer under the necessity of developing structural modifications to survive."

EMOTIONAL HAZARDS IN ANIMALS AND MAN. By Howard S. Liddell. xi + 91 pp. Charles C. Thomas, Springfield, Illinois, 1956. $2.50. Six lectures, delivered in Halifax to psychiatric residents of the Royal Victoria Hospital, present the implications of the study of conditioned reflexes, initiated by Pavlov.

MINNESOTA STUDIES IN THE PHILOSOPHY OF SCIENCE, VOL. I. Edited by H. Feigl and M. Scriven. xiv + 346 pp. University of Minnesota Press, Minneapolis, 1956. $5.00. "The first volume of Studies presents some of the relatively more consolidated research of the Minnesota Center for the Philosophy of Science and its collaborators." Some studies deal with the more general issues (including meaningfulness of theoretical concepts) while others are concerned with more specific topics (psychoanalytic theories, radical behaviorism, validity in psychological tests, characterization of a single personality).

CULTURE AND MENTAL DISORDERS. By Ralph Linton. vii + 139 pp. Charles C. Thomas, Springfield, Illinois, 1956. $4.50. Examines the relationship of psychiatry to cultural anthropology in 3 chapters concerned with Culture and normality, Cultural influences in neurosis and psychosis, and Homosexuality and the hysterias in cultural perspective.

SOCIAL CHARACTERISTICS OF URBAN AND RURAL COMMUNITIES, 1950. By O. D. Duncan and A. J. Reiss, Jr. xviii + 421 pp. John Wiley, New York, 1956. $6.50. The study deals with characteristics of human populations, analyzed in terms of human ecology (social morphology).

EXPERIMENTAL DESIGN. By Walter T. Federer. xix + 544 pp. Macmillan, New

York, 1955. $11.00. An advanced textbook, growing out of the author's experience as professor of biological statistics at the College of Agriculture, Cornell University.

FRANZ JOSEPH GALL, INVENTOR OF PHRENOLOGY AND HIS COLLECTION. By E. H. Ackerknecht and H. V. Vallois. 86 pp. University of Wisconsin Medical School, Madison, 1956. $1.50. Published as No. 1 of *Wisconsin Studies in Medical History*, the monograph deals with Gall's life and personality, his work as a scientist (in anatomy, psychophysiology, and anthropology), and as a social reformer. His collection (busts, masks, skulls, brains) is described and a special study of a few items, including Gall's skull, is presented.

TECHNIQUES IN BLOOD GROUPING. By I. Dunsford and C. C. Bowley. xiv + 250 pp. Charles C. Thomas, Springfield, Illinois, 1955. $4.50. This practical manual describes the principles of blood grouping, presents systematically the laboratory procedures, and provides a glossary of terms and symbols.

JOSEF BROŽEK

University of Minnesota,
Minneapolis, Minnesota.